THE ANALYSIS OF SPOT PRICE STOCHASTICITY IN DEREGULATED WHOLESALE ELECTRICITY MARKETS

A thesis submitted in partial fulfilment of the requirements for the Degree of Doctor of Philosophy in Management Science in the University of Canterbury

by

James P. Tipping

March, 2007

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

ACKNOWLEDGEMENTS

Over the past three years I received many words of encouragement from my friends, family, colleagues and past PhD students, and I would like to acknowledge the love and support that you have all given me. I have loved my time at university, and in particular spending four years teaming up with Paul Stewart both in the office and the indoor cricket arena with the MANGlers!

Thank you to my supervisors, Don McNickle, Grant Read and Deb Chattopadhyay, for the many hours of your time spent advising, assisting, and critiquing, but most importantly in sharing your knowledge and wisdom with me. I have learnt an incredible amount and I owe most of that to you. I have also received many hours of assistance from the staff at CRA International (Asia Pacific) Ltd, as well as generous financial support, and I also received funding from the Tertiary Education Commission. This thesis was examined by Jim Bushnell of the University of California, Berkeley, and Professor Graeme Guthrie of Victoria University, Wellington. Thank you very much for your constructive comments, recommendations and additions to this thesis.

To my wife, Kym, thank you so much for your love and encouragement. Without you, I would have taken twice as long to accomplish half as much!

My biggest thank you goes to my Father in heaven and my Lord and Saviour, Jesus Christ, from whom I receive all strength, wisdom, peace, grace, life and love.

ABSTRACT

Traditionally, time series of wholesale electricity market spot prices have been modelled either by mimicking market operation and equilibrating demand and supply, or by specifying an exogenous process for prices. More recently, a number of hybrid models have been developed, combining the merits of both methods. In this vein, we present an econometric model for daily spot prices in the New Zealand Electricity Market (NZEM) that utilises reservoir management theory to incorporate information on the hydro storage level, a recognised driver of NZEM spot price behaviour. In order to forecast future storage levels and prices, we also construct a model for daily reservoir releases that can be used in conjunction with time series of inflows. This analysis reveals that releases in New Zealand are driven primarily by hydrological factors, as opposed to market conditions. The combined price and storage forecasting model can be applied in a variety of contexts, and offers an alternative perspective to the traditional models of NZEM behaviour. Finally, we calibrate a Cournot model of market behaviour in the National Electricity Market of Australia during daily peak, shoulder and off-peak periods, adding credibility to the future application of such models. The resulting model parameters are, in general, consistent with conventional wisdom. Spot prices from this market are then modelled by combining the output of the analytical model with a stochastic price process.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AR	Autoregression
ARIMA	Autoregressive Integrated Moving Average
ARMA	Autoregressive Moving Average
AUS	Australian
BR	Best-response
CFD	Contract for differences
CML	Constrained maximum likelihood
CMLE	Constrained maximum likelihood estimation
DR	Dynamic Regression
ECNZ	Electricity Corporation of New Zealand
EOY	End of year
EPV	Escribano, Peña and Villaplana
GARCH	General Autoregressive Conditional Heteroskedasticity
GW	Gigawatts
HVDC	High Voltage Direct Current
ISO	Independent System Operator
LRMC	Long-run marginal cost
MA	Moving Average
MAD	Median Absolute Deviation
MC	Marginal Cost
MEL	Meridian Energy Limited
MLE	Maximum likelihood estimation
MOE	Ministry of Energy
MRJD	Mean-reverting jump-diffusion
MW	Megawatts
MWh	Megawatt Hour
MWV	Marginal Water Value

NEM	National Electricity Market
NI	North Island
NSW	New South Wales
NZED	New Zealand Electricity Department
NZEM	New Zealand Electricity Market
OE	On Energy
PCL	Pseudo contract level
PDC	Price duration curve
PDF	Probability density function
PED	Pseudo elasticity of demand
PJM	Pennsylvania-New Jersey-Maryland
QLD	Queensland
RMSE	Root mean squared error
RSL	Relative Storage Level
SA	South Australia
SFE	Supply Function Equilibria
SI	South Island
SIC	Schwartz Information Criterion
SOE	State-owned enterprise
SRMC	Short-run marginal cost
T-CONE	Transmission-constrained Cournot-Nash Equilibrium model
VAR	Vector autoregressive
VIC	Victoria
VoLL	Value of Lost Load
WAB	Water Allocation Board
WVS	Water value surface

1

INTRODUCTION

1.1 Deregulated wholesale electricity markets

In the early 1990s, there began a trend across the world toward deregulating national electricity systems. The extent of the deregulation since then has varied between and within individual countries, but most often it has involved introducing competition between electricity generating companies under a market structure. Central planning has largely been removed, with companies responsible for their own operation of, and investment in, electricity-generating plants.

Electricity systems have always be categorised as the sum of three main parts – generation, transmission and distribution – with the focus of this thesis being largely on the former. In deregulated markets, generating companies offer specific amounts of generation to the wholesale (or spot) market at specific costs per unit. An Independent System Operator (ISO) then equilibrates that aggregate market supply (i.e. all the generation offered) with demand (*load*), setting a spot price where the supply and demand curves intersect and thereby clearing the market. For this thesis, the level of the spot price, and its unique volatility, are the most crucial features of the market operation. All those generators who offered to generate electricity at or below the spot price are dispatched either partially or fully, and they receive the spot price for each unit that they generate. The electricity generated is then transmitted from the point of generation to the end-users around the transmission network.

A major difference between the deregulated and centralised systems is the presence and operation of the wholesale (or spot) market for electricity. In several of the previous regimes, such as New Zealand's, a central authority (i.e. the Government) owned all the generating plants, decided which plants would operate to meet load and determined how much consumers would pay for power[1]. The goal of the system at that time was to meet load at the minimum total cost of generation. The aim now, in deregulated markets, is for the individual companies who own the plants to maximise their profits through generation (or not, as the case may be). Instead of electricity prices being set by a regulator, as in previous regimes, the price is now set by competition in the market. The ISO is responsible for guaranteeing that supply equals demand (at minimum cost, given the offered generation) and determining who generates how much electricity. The spot price of electricity is the all-important variable in the generating companies' profit equation, and modelling its behaviour is crucial for their decision-making.

1.2 Modelling electricity spot market prices

Accurate modelling of spot price behaviour is crucial not just for the strategy decisions of current owners of plant, but also for potential investors calculating expected income from generation. Forecasts of future prices are also important for the pricing of financial derivatives, such as call options for specific amounts of power. Further, due to the high

[1] In the United States, a large proportion of the electricity generating plants were privately-owned and operated, however the price of electricity was still regulated.

risk of the exercising of market power in electricity markets, regulators require models of price behaviour in order to identify and examine periods when such exercising may have occurred.

Due to the relatively recent onset of deregulation around the world, modelling of electricity prices using historic market data has only been undertaken in the past decade, although the volume of literature on the subject has grown rapidly over the past five years. Models can generally be classified as one of two types: top-down, or bottom-up. Top-down models (or time series, statistical or econometric models) do not attempt to model the actual electricity systems themselves, but instead look at series of historic prices and attempt to infer aspects such as price trends and volatility directly from those series. Bottom-up models, in contrast, may include factors such as the marginal costs of generation and generating capacities, transmission constraints, and load, and calculate prices from this information in much the same way as do the actual market dispatch algorithms.

1.3 The aims of our research

As discussed in the following chapters of this thesis, each type of model has strengths and weaknesses. For example, top-down models model volatility due to unforeseen or inexplicable events, some of which cannot be captured in a bottom-up model. However, as they do not (in general) include system factors, top-down models are less equipped to model substantial demand- or supply-side shifts, such as divestiture of generation assets or the increase in price for a particular type of fuel. The overall goal of this thesis is to combine the two types of models in order to utilise the strengths of both in a single framework.

Our aim in completing this research is threefold. Firstly, and most importantly, we want to produce improved fits to historic market data. This adds credibility to any price forecasts made using our models as opposed to others. Secondly, we want to calibrate bottom-up models with existing market data. Before deregulation, the behaviour of hypothetical markets was simulated using bottom-up models. However, since the formation of markets, very little work has been undertaken to analyse how accurate the price behaviour predicted by these models actually is. The calibration also validates the models for forecasting purposes. Thirdly, improved price models which take into account system factors are better tools for testing market designs and for forecasting and analysing market behaviour. We want to produce models that can forecast or backcast price levels and volatility, given various historic or hypothetical situations. This enables the models to find application in a regulatory context, which is becoming ever more important due to the increased potential for the abuse of market power in the electricity sector.

With these aims in mind, this thesis presents models incorporating aspects of both types of models, in the context of two different electricity markets. The New Zealand Electricity Market (NZEM) is a hydro-dominated system, with around 65% of average annual generation produced from hydro generation. Water is therefore the 'fuel' that produces most of the generation in the country, and although water is a free resource, it must be assigned a value in order to balance the risk of running out of water later with the cost of having to use thermal generation now. Despite having such a high proportion of hydro generation, aggregate national hydro storage in New Zealand is very limited, making this balancing calculation very sensitive to the reservoir inflows over just a few weeks. As a result, the level and volatility of prices fluctuate widely depending on the amount of water in the reservoirs. Even though this relationship is accepted as common knowledge in New Zealand, to our knowledge no research exists in the academic literature on incorporating reservoir storage levels into a time series model for NZEM spot prices. We therefore incorporate aspects of hydro-reservoir management theory into a top-down price model in order to capture that relationship.

In order that forecasts can be made using the NZEM price model, we also present a top-down model for New Zealand's reservoir releases. Again, to our knowledge, no such model exists in the academic literature. The model for releases is calibrated using market data, so that it represents how the market has behaved, given the hydrology observed. The release model is used in conjunction with a sequence of historic or synthesised inflows to forecast (or backcast) storage levels, from which price forecasts can be calculated. Applying the combined price and storage simulation model to certain historic situations, using the inflow sequences that were actually observed in those situations, provides interesting insights into how the market could have been expected to behave.

Electricity in Australia's National Electricity Market (NEM) is, in contrast, produced predominantly by coal- and gas-fired thermal generation. As a result, the marginal costs of generation are simpler to estimate than for hydro generation in New Zealand with respect to the uncertainty around the arrival of fuel. Given this information, a relatively simple single-period bottom-up "gaming" model can estimate the price level. Australia's prices are, however, characterised by periods of extreme short-term volatility, which often cannot be accounted for by inputs to a bottom-up model, and a process which accounts for this stochasticity is required. In the latter part of this thesis we present the combination of a calibrated bottom-up model and a stochastic price process to model electricity prices in the NEM. The particular bottom-up model we present, a Cournot model, has been used to synthesise electricity prices in hypothetical markets for many years, however very little research has been undertaken in calibrating the model's parameters with real market data. The research presented in this thesis therefore reinforces the appropriateness of using Cournot models to simulate the operation of electricity markets.

1.4 The structure of this thesis

The remainder of this thesis is structured as follows:

In Chapter 2 we summarize the top-down and bottom-up price models that exist in the academic literature, explaining each of the features inherent in spot price time series for which these models attempt to account. As the primary emphasis in this thesis is on the combination of top-down and bottom-up models, we focus particularly on research that has included physical system information in top-down models. We illustrate the particular relevance of our research in the context of the literature.

Chapter 3 contains background on the New Zealand Electricity Market, from its inception in 1996 to the present day, and details previous studies of NZEM spot prices. We include many references from the media regarding the hydro storage situation in recent years and the impact this has had on prices. With this in mind, we apply a particular top-down model to the series of spot prices from the NZEM to illustrate why the hydro reservoir levels are an essential ingredient to any medium to long term model of that series.

Building on the illustrations in the previous chapter, in Chapter 4 we introduce the marginal water value (MWV) concept from hydro reservoir management theory, and explains its relevance to a model of NZEM spot prices. We demonstrate various relationships between hydro storage levels and spot price levels, and propose a concept we call the Relative Storage Level (RSL) to show the aggregate storage situation relative to the time of year. As the RSL decreases, the risk of running out of water increases, and the price of water (i.e. the MWV) increases to reflect the increased risk. The rise in the MWV (and increased cost of hydro generation) flows through to generators' offers to increase the spot price. Finally, we incorporate the RSL into the same top-down model as in Chapter 3 to improve the overall fit to NZEM spot prices.

In Chapter 5 we extend the price model by incorporating the water value into the stochastic component of the top-down model. The reasoning behind this is that both the occurrence and extent of price volatility increase as the RSL decreases. We find that, as expected, linking certain aspects of the stochastic price process to the MWV does increase the overall fit of the price model, and reflects the intuition behind the model.

As we propose that both the underlying level and the volatility of spot prices in New Zealand is based on the RSL, in order to forecast prices some method of forecasting the RSL is required. The two determinants of future storage levels are future inflows and future releases. While inflow modelling is a well-established area of the hydrology literature, very little (if any) attention has been paid to modelling time series of releases. In Chapter 6 we explore the relationships between the logical drivers of release and actual releases, and compile a simple model to forecast release using the drivers we find to have statistically significant relationships with release. At the conclusion of that chapter, we explain how this model can be combined with a series of inflows and the price model to simulate storage levels and prices in a forecasting or backcasting context.

Using the combined simulation model of Chapter 6, in Chapter 7 we present potential applications to the NZEM. These include: the examination of historic storage and price situations, both prior to and since the start of the market; assessing the possible impact on NZEM spot prices of extreme inflow sequences or imposing constraints on storage levels and/or releases; and calculating a 'long run' price duration curve (as opposed to the price duration curve observed in the relatively short period since the market started). Each of these applications gives interesting results and is useful for generators, investors and regulators alike.

In Chapter 8 we shift our focus from the NZEM to Australia's National Electricity Market. We calibrate an existing bottom-up model using real market data, discussing the intuition behind the resulting parameters. We then fit the variation in prices not explained by the bottom-up model with a stochastic price process from a top-down model, and determine whether or not the combination of the two types of models gives a fit superior to the top-down model on its own. The series we fit for the NEM are split into peak, shoulder and off-peak periods for each day, and we compare the estimated model parameters for each of the three periods.

Finally, in Chapter 9, we discuss the overall findings and impact of our research. There are many areas in which the models presented can be extended, and some of these are discussed, with reference to the ideas raised in the course of our research and the questions we have left unanswered.

2

ELECTRICITY SPOT PRICE MODELLING

2.1 Why model electricity spot prices?

In the past decade, increased deregulation within electricity markets around the world has resulted in competition between the companies responsible for generating power. The reductions in regulation and government price-setting in the market have led to wholesale electricity spot market prices becoming much more volatile. Increased price volatility has resulted in, among other things, electricity market participants facing increased risk, both in terms of the volumes of electricity they can produce and sell, and the prices they will receive for their output. In order for participants to make informed decisions with respect to operations, risk management, and investment, it is therefore vital that they have accurate tools for modelling spot price behaviour.

Commodity price modelling is not a new area of research by any means. However, as explained in later in this chapter, electricity spot prices exhibit a range of characteristics that render the traditional price models unsuitable. Due to the fact that electricity markets around the world are only now maturing, electricity spot price modelling is becoming an important focus of study, and, in recent years, the amount of literature available on spot price behaviour and applications of spot price modelling has increased rapidly. Motivating this research are the increasing availability of price data, and, more importantly, the requirement of all sectors of electricity markets to model prices accurately. However, despite the interest in the field of spot price modelling, there are still many areas of research which have yet to be explored.

Forecasts and models of spot prices are required for many different applications in the operation of electricity markets. For example, in the short run, generating companies have to make decisions regarding unit commitment. They will only want their generators to be dispatched if it is going to be profitable to do so, and, as these decisions are often required hours or days in advance, they require forecasts of future spot prices in order to determine profitability[1].

In the medium term, generating companies whose plants need periodic maintenance require spot price forecasts in order to determine the time to take their plants offline that will have the least impact on their profit levels. In the longer term, potential investors in new or existing power plants also need forecasts of spot prices in order to determine the potential profitability of (and return on) their investment. Many other industries use and pay for electricity as an important input in their operations, and they also require forecasts of spot prices in order to determine their own profitability. In many markets around the world, these users (and other speculators) are able to purchase contracts for electricity at a fixed price over a specified time period. The valuation of such financial derivatives

[1] Also, as discussed more thoroughly later in this thesis, some generators have to operate unprofitably in some hours in order to be able to operate profitably in others. See Guthrie and Videbeck (2002b) for further discussion on the short-term requirements of spot price forecasting.

requires estimation of both the likely levels and volatility of spot prices in order to determine what that fixed price should be, as well as a fair price for the contract itself[2].

Aside from generators, investors and consumers, a fourth group of users, regulatory bodies, also requires models of spot prices in order to test hypotheses regarding market behaviour. Due to unavoidable characteristics such as market segmentation and fragmentation due to transmission constraints, electricity markets have in recent times been the subject of scrutiny regarding the abuse of market power[3]. In order to determine whether or not markets are behaving competitively, a common test is to see how far above competitive levels prices are, or have the potential to be. A key requirement for such tests is a model that is able to mimic market price behaviour accurately. Regulators are then able to compare market behaviour in competitive situations with behaviour in other situations, both hypothetical and observed. The outcomes of such comparisons have implications for all parties with an interest in electricity markets.

The remainder of this chapter is structured as follows. Section 2.2 details the features of electricity spot price time series that make them unique, and render the traditional models unsuitable. These characteristics influence greatly the types of models that have been (and must be) developed for forecasting their behaviour. Section 2.3 then provides a detailed overview of the types of models that have been developed for the purpose of modelling electricity spot prices. This section forms the majority of the literature review in this thesis; however, as detailed in the introduction to that section, further reviews of the relevant literature are given in the following chapter, and also in subsequent chapters.

---

2 One type of contract used in many electricity markets is the forward contract. These contracts specify a fixed price (the forward price) at which electricity will be purchased sometime in the future. The forward price is based on the expected level of the spot price at that time, and the forward premium, or the amount that the purchaser of the contract is willing to pay (or the seller is willing to concede) to guarantee the price at which they can purchase electricity, is defined as "the difference between the expected spot price and the forward price" (Longstaff and Wang, 2004).

3 Market power in the electricity market setting is broadly defined as the ability of a firm to offer its generation for sale at a price above its marginal cost of generation, yet still be dispatched.

2.2 Characteristics of electricity spot price time series

Some knowledge of the operation and clearing of an electricity market is required in order to understand the factors that influence spot price behaviour. A brief description of the fundamental concepts is provided in Appendix A.

In the simplified form detailed in that appendix, short run electricity markets are really no more complicated than markets for many other goods. But complications arise for several reasons: spot prices vary at different points in the same network due to transmission losses and constraints; electricity flows follow the laws of physics and not the wants of man; and electricity cannot be bought in one period and stored for later use[4] to enable arbitrage between time periods. Thus, unlike in other markets, electricity, the basic commodity for which the market-clearing price is set, is a different commodity at every single market clearance, and is subject to different supply as well as demand conditions at every hour of every day. The need for a constant supply-demand balance leads to highly fluctuating prices in the short run. As a result of each of these factors, the combinations of characteristics exhibited by electricity prices are different to the prices of all other commodities.

2.2.1 Seasonality

Electricity spot prices in deregulated markets are set through balancing the supply of power from generators with the demand for power by consumers (commonly referred to

[4] Effectively, any water in storage reservoirs is electricity waiting to be generated, but electricity itself cannot be stored in any appreciable scale. Electricity can be stored by consumers in batteries, but this is not a feature of any major electricity systems at this point in time. On the supply side, pumped hydro capacity (where water is pumped back up to the top of a hydro dam after it has been used to generate power) does exist in several markets outside of New Zealand, but each of these storage options involve substantial losses and costs.

as the system *load*). Spot prices are likely to exhibit strong seasonality, on account of both supply and load displaying periodicity.

Spot prices typically exhibit patterns over three different lengths of time. Two of these, intra-day and intra-week patterns, are driven primarily by load, whereas annual patterns are driven by both load and supply. Load varies markedly, according to the time of day (i.e. peak or off-peak), the day of the week (i.e. weekday or weekend) and the time of year (i.e. the season – hot, cold or mild). The intra-day pattern is evident in a series of New Zealand spot prices over the course of a single day, illustrated in Figure 2.1 below, which shows the final half-hourly spot market prices for Thursday 26 June 2003 for the Haywards node in Wellington[5]. It has a clear morning peak, when people are waking up, turning on heating, showering and possibly cooking breakfast, a higher general level during the day on account of commercial activity increasing load, and an evening peak when people prepare their main meal and use other appliances. The difference in price behaviour on a weekday (when people rise early for work and require electricity for commercial activity) compared with a weekend is also understandable[6].

[5] The New Zealand price data was sourced from the New Zealand Electricity Market website, http://www.nzelectricity.co.nz/

[6] Li and Flynn (2004a) offer a more thorough discussion and empirical examination of the intra-day patterns in spot prices.

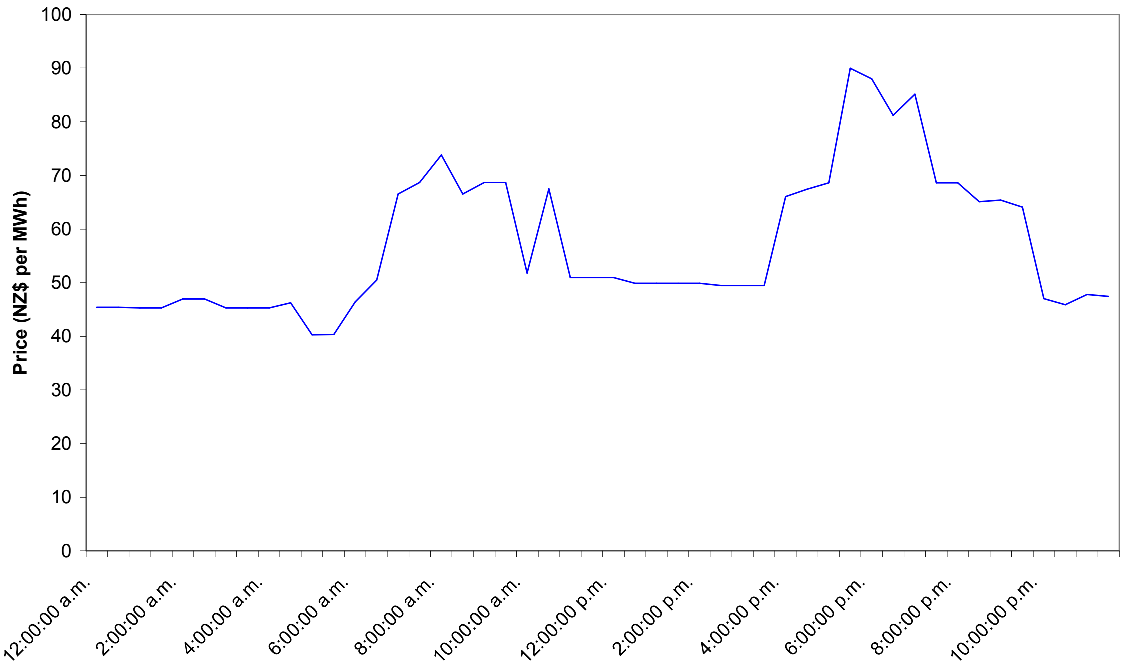

Figure 2.1: New Zealand half-hourly spot prices at the Haywards node for 26 June 2003

Load also varies depending on the time of year. In many European and North American countries, as well as in New Zealand, load is greatest in the winter when electricity is required for heating. In other countries, such as northern parts of Australia, load is greatest in the summer, due to the demand for air conditioning. Some Asian countries, among others, experience two distinct peaks in load during the year.

Supply varies throughout the week and year depending on the availability and cost of generation. This variation is driven by the strategy of generating companies (including maintenance scheduling), the cost of fuel, and most importantly in hydro-dominated markets such as New Zealand's, the amount of water captured by hydro reservoirs. Generating companies may schedule maintenance to occur when their units are not normally required, to minimise the impact on total costs, or, strategically, when the outage will have a large effect on the prices captured by the remainder of their portfolio.

The cost of fuel for thermal generators is unlikely to vary too greatly with seasons, due to the fact that companies will secure the majority of their fuel through long-term contracts.

However, as discussed further in Chapter 4, the cost and availability of hydro generation can vary dramatically over the course of a year. The colder the weather in winter, the more precipitation falls as snow instead of rain, and thus this water cannot be captured by the hydro reservoirs until the temperature warms in spring and the snow melts. The increase in temperature coincides with a decrease in the demand for electricity, and the reservoirs can then be refilled. The relationships between temperature and load, and temperature and supply, are therefore linked, and intensify the impact each has on spot prices. Water can be most scarce when it is most required, often stretching the balance between load and hydro supply during winter.

2.2.2 Price-dependent volatility and volatility clustering

As explained in Appendix A, the shape of the market offer stack tends to be relatively flat for lower levels of load, but steeper as more expensive capacity is offered. This convexity in the shape of the offer stack leads to another interesting phenomenon exhibited by electricity prices, price-dependent volatility[7]. When load is low (and the market-clearing price is low), small fluctuations in load are unlikely to change the market-clearing price significantly, as the offer stack is virtually flat for low load levels. However, for higher levels of load, the offer stack steps get steeper, and small shifts in demand may lead to the market-clearing equilibrium lying on a much higher or lower step than previously. (Also, as suggested in Appendix A, the higher steps of the offer stack may also have been raised artificially as a means of strategic gaming.) This leads to prices being more volatile when they are higher, and also volatility occurring in clusters in periods when the load is high. Figure 2.2 below illustrates how two identically-sized demand shifts at different levels of load (D_1 to D_1' and D_2 to D_2') can have two very different effects on price (p_1 to p_1' and p_2 to p_2'):

[7] Strictly-speaking the volatility is dependent on the load, rather than the price. However, if the only information available was price data, the volatility in prices would appear to depend on the level of prices.

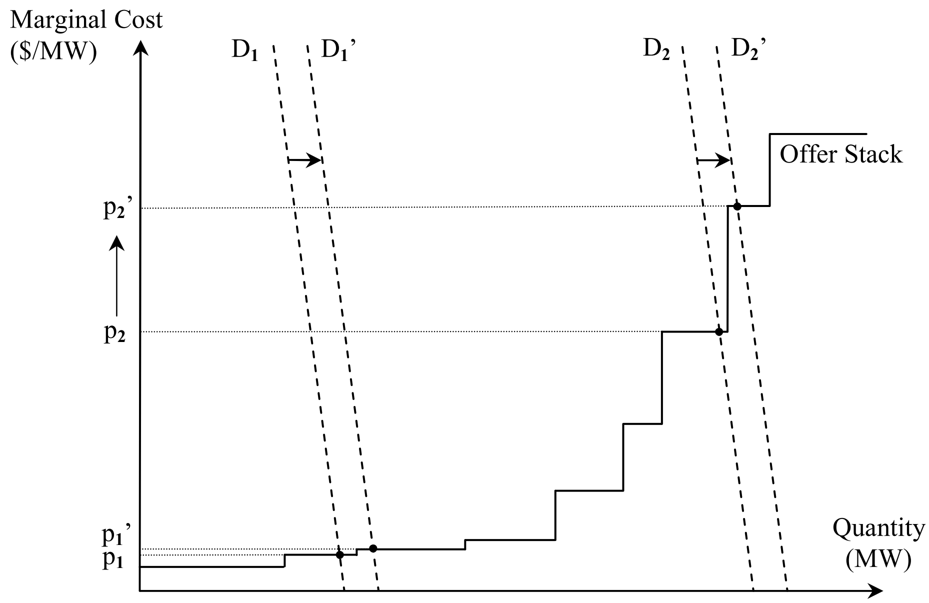

Figure 2.2: Supply and demand curves illustrating price-dependent volatility

The shape of the market offer stack also leads to a phenomenon Knittel and Roberts (2005)[8] label the "inverse leverage effect". In most financial markets, a decrease in the price of an asset leads to an increase in the volatility of that asset's price, which is referred to as the "leverage effect" (Black, 1976, as cited in Wu, 2001). However, the opposite occurs in electricity markets, due to the convexity of the supply curve. A decrease in the load (and price) usually results in the marginal generator being located on a flatter section of the supply curve, thus reducing potential volatility, as shown in Figure 2.2. However, an increase in the price will have the opposite effect on the location of the marginal generator, thereby increasing potential price volatility.

2.2.3 Extreme jumps and spikes in prices

It is important to note that large fluctuations in spot prices can also occur at much lower levels of load. For example, forced outages of generating capacity can truncate the offer stack by removing one or more steps. If one or two of the lower-cost steps were suddenly

[8] The paper by Knittel and Roberts was released firstly as a working paper in 2001. At that time, it was one of the first and most important papers on spot price modelling. The paper was not published until 2005.

removed from the offer stack in Figure 2.2, the effect on prices would be large, both in periods of low and high load. Because electricity cannot be stored in any appreciable scale, there can be no stockpile of electricity to smooth out such shocks in supply and demand, and any shortfall in supply must be met instantaneously with extra generation. The extra units may have a marginal cost many times greater than the cost of the unit they are replacing, resulting in what is referred to as a "jump" in the spot price.

Jumps in prices can arise for any number of reasons, but usually occur when a generator suddenly goes offline, demand increases, or transmission between points in the market becomes constrained. However it is when more than one of these factors occurs at once (such as when limited supply coincides with high load) that prices are at their most volatile. Spot prices can increase in value by a factor of several hundred within one clearance of the market. While the price often decreases by the next time the market is cleared (once the combination of circumstances has been corrected), the price can remain high for several periods afterwards. These extreme rises and subsequent falls in prices are known as spikes, and are a key component of electricity spot price time series, making spot price series much harder to model

At times when the electricity system is not transmission-constrained and there is a large amount of excess capacity, the spot price is relatively well-behaved and can be modelled by any number of financial or statistical time series models designed for markets in which spikes are not common. Fitting and forecasting these spikes is the key to modelling electricity spot price time series however, as they have serious implications for market participants. Investors in peaking plants, which have relatively high marginal costs of production and only operate when the price is extremely high, require estimates of how many periods per year their generators will operate, and thus how great their return on investment would be. The same applies to valuing any other generating asset, regardless of its marginal cost of generation. Forecasting the intensity and likely size of price spikes is also vital for calculating the value of contracts for risk management.

As an illustration of the occurrence of price spikes, Figure 2.3 below shows the final New Zealand half-hourly spot market prices for October 2002 for the Haywards node. Note that the price generally stays around some base level during the whole month; however, there are at least three significant price spikes during this period[9].

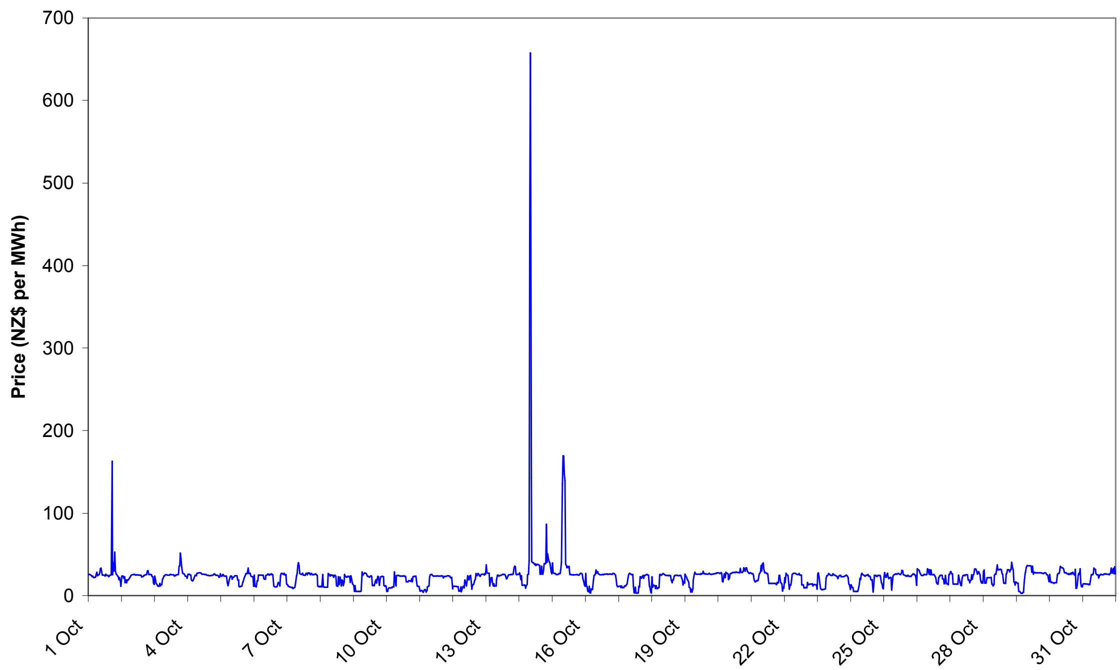

Figure 2.3: New Zealand half-hourly spot prices at the Haywards node for October 2002

Linked with price spikes is another feature that electricity price time series do not have in common with prices of other assets, price caps. In many electricity markets, the spot price is unable to rise higher than a pre-determined value, referred to as the Value of Lost Load (VoLL). This value is calculated as the cost to the system of not being able to satisfy the entire load with supply. In Australia, VoLL is \$10,000/MWh[10]. The price is capped as, theoretically, at times when demand exceeds supply the price would be infinitely high.

[9] See Videbeck (2004) for more examples of price spikes in New Zealand.

[10] http://www.aemc.gov.au/rules.php

2.2.4 Negative prices

While prices are capped from above in most electricity markets, there is often no lower limit on them and, in several periods each year in Australia at least, they take values below zero[11]. Some generators have high start-up and shut-down costs and are slow to ramp up, therefore in order to ensure that they will be dispatched in peak periods of the day, to take advantage of high spot prices, they must also ensure that they are generating in off-peak periods. If there is enough competition between such generators to be dispatched, they may be offered into the market at negative prices and the market may be cleared at a price below zero.

2.2.5 Mean-reversion

Formally, the change in price from period t to period t+1 for a mean-reverting price process (in the discrete case) is the sum of two components:

$$p_{t+1} - p_t = \alpha(\mu - p_t) + \varepsilon_{t+1}$$

where μ is the mean-reversion level or long-run equilibrium price, α is the mean-reversion rate, and ε_{t+1} is the random shock to the price between period t and t+1. The $\alpha(\mu - p_t)$ term is the mean-reversion component. The mean-reversion rate, α, determines how quickly the price will return to its long-run equilibrium price after a price shock, and is restricted to being positive. The greater the value of α, the greater the weight placed on the difference between the long-run equilibrium price and last period's price, and the more quickly the price will revert to its long-run level.

As is clearly illustrated in Figure 2.3, spot prices tend to hover around the same underlying level for much of the time, fluctuating occasionally away from that level but returning in due course. That underlying level is ultimately determined by the marginal

[11] The Australian electricity market does actually have a price floor, which is set at -VoLL.

cost (MC) of generation. Theoretically, in a competitive market companies will offer their generation to the market at their MC[12]. If generation is offered at prices above that level for sustained periods, companies will not be dispatched as often as is profitable, and they will be aware that their actions are likely to attract the interest of market regulators. Offering at prices below that level for sustained periods will be unprofitable. Therefore, the underlying or average level of electricity prices will be determined at the point where load intersects the aggregate market MC curve when transmission is unconstrained and all installed generating capacity is available.

Of course, as mentioned in the sections above, factors such as transmission and plant outages will cause prices to fluctuate away from and around that mean level, even if all generation is offered at MC. Seasonality in the load also flows through to the mean level, as the demand curve intersects the MC curve at higher or lower levels. However, as transmission constraints and generation outages are short-term effects, over time the price time series will always revert to its underlying or mean level. For this reason, electricity prices are said to be mean-reverting.

The economic intuition described above, suggesting the specific characteristics of mean reversion, seasonality, price-dependent volatility and occasional positive and negative spikes, is well supported by empirical research. For example, after conducting a graphical analysis of price data for the markets in California and in the USA, UK, Norway and Victoria, Johnson and Barz (1999) concluded that electricity price series do indeed exhibit each of these characteristics. Kaminski (1997) stated that electricity prices do not behave like the prices of other commodities merely with higher levels of volatility, but instead exhibit a tendency toward jumping suddenly upwards from some floor level, which itself varies with time. This behaviour results from the unique physical

[12] The MC of electricity generation varies depending on the time frame over which it is defined. In the short run, the MC will include just the direct cost of generation, such as the fuel cost, but may include costs that take into account start-up and shut-down costs that need to be recovered. In the long run, extra costs will also be included, such as the cost of constructing the generating plant and the costs of any capital borrowed to finance this construction.

characteristics of electricity and the market operation described above, and is the cause of the leptokurtic distribution[13] of most markets' power prices.

Hundreds of other studies of electricity market prices have been conducted around the world in recent years. Many, if not all of them, begin by describing the characteristics of electricity price time series and the market fundamentals which cause these characteristics. Particularly sound descriptions are provided by Geman and Roncoroni (2006) and Bunn (2004). The actual modelling techniques employed vary widely, with many of these techniques described in the following section.

2.3 Electricity spot price modelling

Green (2003) summarises the questions commonly asked with regard to how best to model spot prices:

> "Should we be following the approach of the finance literature, which treats the price of electricity as a stochastic variable and concentrates on studying its properties in terms of volatility, jumps, and mean reversion? Or should we concentrate on the fact that the short-term price for every period is set by some intersection of demand and supply, and study the interaction of these factors with the market rules? I would be uneasy if the stochastic approach is taken to imply that we cannot understand the out-turn values for each period's price. However, it may be that explaining every price in turn is too cumbersome, and randomness should be taken as a shortcut for "things we could explain, but don't have time for in this application."

[13] A leptokurtic distribution is symmetrical in shape, similar to a normal distribution, but the centre peak is much higher, there are fewer observations in the shoulders and more in the tails.

The tools currently developed for the purpose of modelling and forecasting electricity prices can be broadly categorised into Green's two groups of models, which Davison, Anderson, Marcus and Anderson (2002) classify as "bottom-up" and "top-down" models. Descriptions and examples of both of these types of models are provided in more detail in the following two sub-sections. However, in line with Green's questions, the research presented in this thesis concentrates on neither one type nor the other. Instead, as mentioned in the introductory chapter, we focus on applications of combinations of the two types of model. These applications will therefore form the majority of the literature review in this chapter.

The broad characteristics of the two types of models are illustrated in Figure 2.4. Bottom-up models utilise the theory of electricity markets and equilibrate demand and supply to calculate a price, whereas top-down models specify an exogenous process for prices, the parameters for which are estimated entirely from the series of prices itself. The characteristics of each approach are detailed further in the remainder of this chapter, however Figure 2.4 should be referred to throughout.

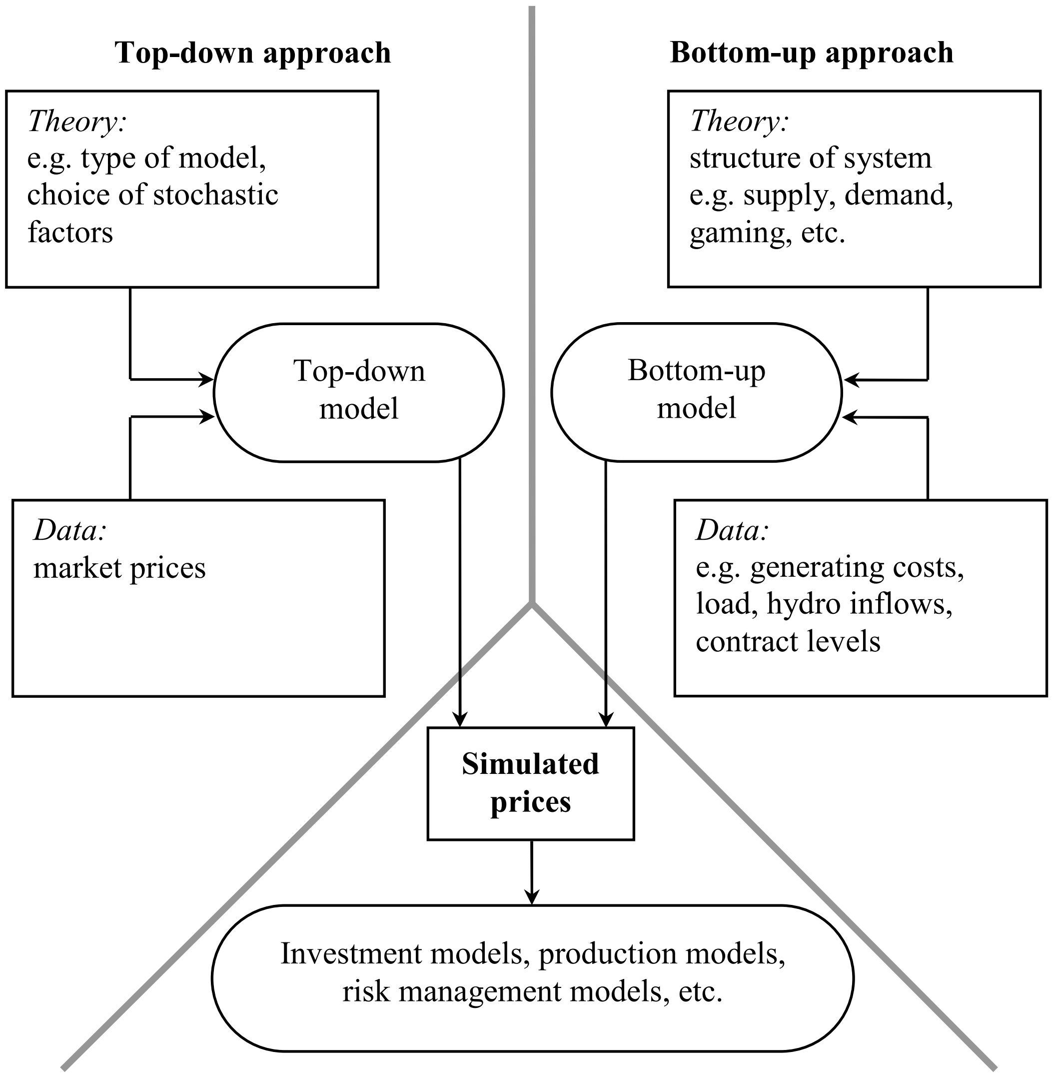

Figure 2.4: Two approaches for transforming data and theory into simulated prices, which can then be used as input to decision models (Adapted from Figure 1 in Fleten and Lemming, 2003)

2.3.1 Bottom-up models

Bottom-up models estimate prices in much the same way as the market-clearing process of an electricity market actually occurs, which is described in Appendix A. In their simplest form, they firstly construct the market offer stack, using information on all the generating capacities and marginal costs of each generating company. They then calculate the spot price by finding the intersection of the market demand curve with the market offer stack, with more sophisticated models taking into account locational and transmission issues as well. Bottom-up models are, understandably, very data intensive.

Bottom-up forecasts of the price over many time periods obviously require forecasts of all the supply- and demand-side information for those periods. To do so, the model may calculate generation levels and prices for each period independently and repeatedly, in which case it is referred to as a static equilibrium model. However, power systems that are constrained inter-temporally, such as thermal systems with limited fuel stockpiles or hydro systems with storage reservoirs, require more complex multi-period optimisation over the entire period for which price forecasts are required. This is because the optimal allocation of fuel resources (among other things) also needs to be determined across the period in question. In these cases, such a model can be called a dynamic equilibrium model[14]. Extending the literature on these types of models is not the aim of this thesis; however references for some of the research into the operation of mixed hydro-thermal systems are given in Chapter 4.

2.3.1.1 Behavioural / game-theoretic bottom-up models

Every bottom-up model requires some assumption regarding how each generating company will offer its generation into the market. For example, the techniques described above in Section 2.3.1 assume that each unit of generation will be offered into the market at its marginal generating cost, and not above or below. They do not take into account other aspects, such as forward contracts, that may influence a generating company's bidding strategy. In essence, they assume that generating companies act as perfect competitors, whereas many studies claim instead that generating companies in fact exercise market power[15]. However, where these models are useful is in calculating a competitive benchmark for prices, to which actual prices can be compared (see Borenstein, Bushnell and Wolak, 2000; and Bushnell & Saravia, 2002).

---

[14] These models can be deterministic, if factors such as fuel availability are known with certainty ahead of time, or stochastic, if many different scenarios (with different probabilities of occurring) are considered.

[15] See Pirrong and Jermakyan (2001) for references to these studies.

Perhaps more common in the recent academic literature are models predicting market behaviour as a result of companies following some pre-specified offering strategy, such as the Cournot or Bertrand competition models, or the more realistic but computationally more complex Supply Function Equilibrium model. However, as mentioned in Chapter 8 of this thesis, while the application of these models for predicting market behaviour in hypothetical situations is widespread[16], very little work has been done in assessing just how accurate these models are ex post, which is one of the foci of this thesis.

In general, the strength of bottom-up models is their ability to look forward and predict market behaviour in hypothetical future situations. However, the fact that very little work has been undertaken in identifying their ability to model current behaviour is concerning, and is the motivation for the work presented in Chapter 8. While bottom-up models are used in practically every market around the world, very few applications of such models, calibrated with and tested against real market data, exist in the spot price modelling literature.

One example of the application of a bottom-up model to estimate prices is the study of Bessimbinder and Lemmon (2002), who equilibrate demand and supply to model prices from the Pennsylvania-New Jersey-Maryland (PJM) and Californian markets in the United States from 1997 to 2000. In their model, each generating company chooses its output for the period to maximise its profits, given its total cost curve (with marginal costs increasing in output) and the market demand curve. The total output of all the companies must equal the load in that period, and the equilibrium market price is calculated as a function of the ratio of the system load to the number of generating companies. In this way, only the load varies by period; the number of companies is included as a proxy for total generating capacity, and remains constant each period.

---

[16] Such hypothetical situations may include the formation of a new market, the division of a large generating company into several smaller companies, the entrant of a new competitor into a market, a fuel shortage or surcharge, or the imposition of forward contracts on generating companies.

Another recent example of the application of a bottom-up model to pricing electricity contracts is that of Fleten and Lemming (2003). They forecast three years of spot prices for the Nordic region using a bottom-up model, and then adjust their forecasted spot prices so that they are more in line with current futures prices for the same period of time.

Further examples of static equilibrium models can be found in Green and Newberry (1992), Borenstein and Bushnell (1999) and Anderson and Philpott (2002). Studies using dynamic equilibrium models (i.e. calculating prices in more than one period simultaneously) include Garcia and Arbelaez (2002) and Garcia, Campos-Nañez and Reitzes (2005). However, while some of these studies' input parameters were estimated using real data, none of the models were calibrated with observed price data to improve their performance in modelling prices.

2.3.2 Top-down models

In contrast to bottom-up models, the strength of top-down models is in looking backwards at past price behaviour. The parameters of these models are estimated directly from observed market data, and they are therefore able to draw inferences regarding such things as seasonal patterns in the prices, structural changes and patterns of volatility. As discussed further in Chapter 8, in general bottom-up models do not have the same capability as top-down models to forecast volatility in prices. Whereas bottom-up models require a great deal of physical system information in order to estimate a price for electricity, top-down models specify a process for the price that may be completely independent of the underlying physical system state variables, and usually only rely on past realisations of the price. Therefore, their advantage is that often the only input required for a top-down model of spot prices is an observed time series of the actual prices.

Traditionally, bottom-up modelling has been the only way in which electricity spot prices could be estimated, due to the fact that no time series of observed market prices existed. However, in many sectors, top-down modelling of prices is now a much more prevalent form of price modelling than bottom-up, and the amount of literature on top-down electricity price modelling has increased exponentially in the past decade as more historic price data has become available. Several of the major methods employed are broadly defined below, with examples of each provided.

2.3.2.1 Econometric models

The widest and most extensive group of top-down models in the literature are econometric or time series models. They involve the application (and extension) of models previously developed for and applied to time series of other asset prices. Knittel and Roberts (2005, but originally 2001) fit various financial models of asset price processes to hourly electricity prices from the Californian market, and conclude that none of the standard models they try are suitable for fitting to electricity prices. Since their study, new econometric models have been developed which attempt to account for all the features inherent in electricity price time series.

The most commonly component of such models is mean-reversion. As mentioned in Section 2.2.5, electricity prices may fluctuate from day to day, but in general in time they will return to an underlying level determined ultimately by the MC. Two of the first authors to propose a mean-reverting model for spot prices were Lucia and Schwartz (2002) in their study of Norwegian electricity derivative prices. They propose that prices could be decomposed as the sum of two components, a deterministic component, $f(t)$, and a stochastic component, X_t:

$$P_t = f(t) + X_t$$

In their model, the deterministic component represents the underlying mean level of prices, and accounts for seasonality in the prices. The stochastic component accounts for decaying movement away from that level, in the form of autoregression (AR). Shocks to the price series have a lasting effect on the level of prices which decays over time. Therefore, while prices may fluctuate from day to day in the model, they fluctuate *around* an underlying level modelled by the deterministic component. Other studies proposing mean-reverting models include Johnson and Barz (1999) and de Jong and Huisman (2002).

In time, it became evident that price spikes were a key feature in electricity price time series, and models incorporating jump processes had to be developed. Like other shocks to electricity prices, the effects of jumps decay over time, in some cases very rapidly. A class of models referred to as mean-reverting jump diffusion (MRJD) models has become the most popular class of models applied to electricity prices in recent years, with several classes of jump processes proposed. The MRJD model of Escribano, Peña and Villaplana (2002) is described in detail in the following chapter, however other examples of the application of MRJD models include Deng (2000b), Atkins and Chen (2002), Eydeland and Wolyniec (2003), Villaplana (2003), Natarajan (2003), Goto and Karolyi (2004), Borovkova and Permana (2004), Chan and Grey (2005), Knittel and Roberts (2005), and Cartea and Figueroa (2005).

Various processes accounting for heteroskedasticity[17] in electricity prices (as a result of price-dependent volatility and clustering in the volatility) have also been incorporated into these econometric price models. Most of these have been based on AR methods developed for modelling other commodity prices. See Robinson (2000) and Duffie, Gray and Hoang (1999) for examples.

2.3.2.2 Regime-switching models

In its simplest form, a regime-switching model states that the price may take one of two values on day t: p_1 if the system is in state 1 on day t, and p_2 if the system is in state 2. If the system is in state 1, there is a certain probability Pr_{12} of it switching to state 2 each day (and probability (1- Pr_{12}) of staying in state 1), and vice versa if it is in state 2.

---

[17] Heteroskedasticity is the presence of non-constant variation in the residuals of a model.

In applications to electricity markets, researchers have justified the use of these models by stating that “typically, the price process is divided into two regimes: one for the ‘normal’ process, one for the ‘spikes’” (de Jong, 2005). Characteristics of the price series, especially the level and volatility of the prices, vary depending on the regime; prices will be high and volatile in the spike regime, and lower and more stable in the normal regime. These regimes may be driven by the underlying states of demand and supply, as described in some of the studies referred to in Section 2.3.3.

One of the earliest studies in modelling time series of electricity prices with a regime-switching model was undertaken by Ethier and Mount (1998). Since then, studies have been completed by Deng (2000a), de Jong and Huisman (2002), Huisman and Mehieu (2003), Bierbrauer, Trück and Weron (2004), Haldrup and Neilsen (2004), and de Jong (2005).

2.3.3 Combination/hybrid models

As explained earlier, the strength of bottom-up models is their ability to predict market behaviour in hypothetical situations for which no historic data is yet available. Once calibrated, they can predict the likely effects to prices of changes in demand and/or supply conditions. Top-down models, in contrast, require historic price data as their only input, and are able to identify patterns in (and processes for) past price behaviour. Obviously each type has advantages and disadvantages in this regard, therefore an increasing number of hybrid models are being developed, taking advantage of the merits of both types of models.

To our knowledge, there are no bottom-up models which incorporate exogenous stochastic price processes. However, there are a growing number of top-down models incorporating information regarding supply and demand that previously was only used in bottom-up modelling. These hybrid models are the major focus of this thesis, and a summary of each of them is given in the following sections. For the purposes of our research, the factors of interest in each study are not necessarily the modelling methods employed, but are instead the types of information that have been incorporated.

2.3.3.1 Top-down models incorporating demand-side factors

The most common demand-side variable incorporated into top-down models is the aggregate system load. The rationale behind including this variable is very intuitive – with all other factors (especially the aggregate market offer stack) held constant, and assuming an unconstrained transmission network, the greater (lower) the load, the greater (lower) the price will be. This is why, in general, prices are higher in peak load periods during the day than in off-peak periods during the night.

Vucetic, Tomsovic and Obradovic (2001) completed one of the first studies incorporating load as an explanatory variable in a model for price. They regress the spot price on load to examine the relationship between the two variables and find that the relationship is significant, but more importantly they are able to identify changes in regimes based on changes in the nature of the relationship.

Nogales, Contreras, Conejo and Espínola (2002) model price as a dynamic regression, with the explanatory variables including past observations of the price, and current and past observations of load. Surprisingly, they found that in the weeks for which they forecasted prices, load was not necessary as an explanatory variable for forecasting purposes. Obviously there was enough information in previous price observations without having to glean any residual information from the load. Weron and Misiorek (2005) completed a similar study to Nogales et al, and also find that including the current load as an explanatory variable does not improve forecasting performance.

Popova (2004) examines the interaction of load and price in different zones of the PJM market. She models price in each zone i at time t as a function of variables including the load and temperature in zone i at time t, and finds that while load is a significant explanatory variable at the 5% level, temperature is not. Longstaff and Wang (2004) also include weather variables (temperature and wind-speed) as well as the expected load and load volatility in their regression models of forward premia for the PJM market. They find that in general, the expected load is significantly positively related to the forward premium, but load volatility is a significant variable in only around half of the regression equations, and increased volatility has a negative impact on forward premia. The analysis of forward premia is outside the scope of our research, however.

No firm conclusions regarding the suitability of demand-side explanatory variables can be drawn from these studies, a result reinforced by Li and Flynn in their 2004a study, which is described in the following chapter. While these variables do have intuitive justification for being included, the results of including them in models of price (without also including supply-side factors) have been mixed.

2.3.3.2 Top-down models incorporating supply-side factors

Supply-side factors also have a strong justification for being included in top-down models for price. In the same way as a shift in load will induce a shift in the price, if load is held constant and the market offer stack is shifted up (down), the price will increase (decrease). Such a shift may be induced by a change in fuel prices (thereby changing the marginal cost of generation), planned and unplanned unit outages (which truncate the market offer stack), and the overall availability and maximum capacity of generating units.

As an example of including marginal cost information in a model for prices, Guirguis and Felder (2004) compare the forecasting performance of four different top-down specifications for the spot price in the PJM market, each including lags of the prices of natural gas and oil, with the rationale that these are “two fuels used by marginal generation units”. As such, the cost of these fuels would directly influence the marginal cost of generation (and presumably the offer price) of these units. They find that the parameter estimates for oil prices were not significant, however those for the previous day’s gas price were. This is unsurprising, as most power stations are likely to have large stockpiles of oil, but only a limited supply of gas, and may be calculating their SRMC on the basis of the replacement cost of gas.

As explained by Westergaard, Mullon, Sise and McCord (1996) and many other studies, the amount of water available in storage has a strong influence on hydro generators' behaviour, and in particular the prices at which they offer to generate power. This influence will be greater in those markets that have a high proportion of hydro generation, such as in Norway, Chile, Brazil and New Zealand. Several models of electricity prices have been developed which incorporate of hydrological information. These are detailed in Chapter 4, which focuses on the impact of hydro storage levels on spot prices in New Zealand.

2.3.3.3 Top-down models incorporating both demand- and supply-side factors

Far more numerous than either of the previously-described two types of hybrid price models are models that incorporate information regarding both the demand and supply of electricity. Several of these models include supply and demand information independently of each other, while others utilise the fact that it is the interaction between supply and demand that sets the price.

Possibly the most widely referenced authors in the literature on electricity price modelling incorporating physical factors are Pirrong and Jermakyan (1999, 2001). Their work on electricity derivative pricing develops price models based solely on two variables: the fuel price (as a proxy for thermal generators' marginal cost), and load. They developed two different models which specify the price at time t (P_t) as the product of a function of the marginal generator's fuel price (g_t) and a function of the load (q_t). Their first model (1999) specifies price as the product of the fuel price raised to a power and an exponential function of load, as shown below.

$$P(q_t, g_t, t) = g_t^{\gamma} \exp\left[\alpha q_t^2 + c(t)\right]$$

In their model, *c(t)* is a deterministic seasonal function that accounts for seasonality in both demand and supply, that "shifts the pricing function up or down over time". Their second model (2001) is similar, except the demand-side term is a function of the natural log of load rather than an exponential function of load. They use these functions as a result of their (intuitive) suggestion that "the price function is increasing and convex in [load]".

In a similar vein, Villaplana (2005) also models price as the product of supply- and demand-side information, specifying price at time t (P_t) as the product of available generating capacity (C_t) and an exponential function of load (D_t):

$$P_t = C_t^{\gamma} \beta \exp[\alpha D_t]$$

However, instead of simply treating load and the available capacity as exogenous variables, Villaplana also goes one step further than Pirrong and Jermakyan. He models both as independent processes, in much the same way as many top-down price models model price, taking into account features such as seasonality[18]. This enables him to forecast spot and forward prices in the future.

Skantze, Gubina and Ilic (2000) also develop models for both the load at time t (q_t) and the position of the supply curve on the quantity axis (b_t), combining them in the following simple formula for the price:

$$P_t = \exp[\alpha q_t + b_t]$$

They assume that the supply curve has a fixed shape (estimated using historic data on fuel prices, maintenance schedules, strategic bidding, etc.) but that its position changes over

[18] In a similar study, Barlow (2002) also proposes a diffusion process for load. His price function, however, is simply a convex function of load with no supply-side variables.

time. Their model for load incorporates factors such as seasonality, uncertainty, mean-reversion and stochastic growth.

As explained in Section 2.2.3, positive price jumps may occur at times when the electricity system is at risk of becoming capacity constrained, either from an increase in load or a decrease in available generating capacity, or both. At such times, one or more expensive generators must be dispatched, causing prices to increase significantly. Incorporating some measure of excess capacity or "freeboard" (i.e. available capacity minus system load, or load divided by capacity) in a model for prices is a logical way to account for the occurrence and size of price spikes. Such measures are now common in spot price models, several of which are detailed below.

Davison et al. (2002) propose a model that includes such a feature. They present a multi-modal pricing model that samples daily average prices from a probability density function with two normally-distributed peaks – one for low prices and one for very high prices. The price model chooses between a low price and a high price based on a function of the ratio of load to available capacity. When this ratio is below approximately 0.7, there is only a very low probability (<0.01%) of a price being drawn from the high price distribution[19]. However, as the ratio approaches 1, the probability of a high price increases towards 100%.

Burger, Klar, Müller and Schindlmayr (2004) calculate a similar ratio in their study of European electricity prices. They model price with both long- and short-term stochastic processes, and include an exponential function of the *adjusted* load, whereby load is divided by the "average relative availability" of power plants (represented by a percentage of maximum availability). For example, they state that in Germany in January all power plants are available to generate, therefore the adjusted load and load will be

[19] Birnbaum, Del Aguila, Dominguez and Lekander (2002, as cited in Villaplana, 2005, p. 28) find that price spikes occur when this ratio is as low as 0.75, however the probability of them occurring when the ratio is any lower is very small.

equal. However, when more generating units are taken offline in the summer months, the load will be adjusted upwards.

In their regression of the Lerner Index[20] on several explanatory variables for the British electricity market, Evans and Green (2005) also find the demand/supply ratio to be a highly significant variable. This suggests that the tighter the balance between demand and supply, the greater the opportunity to exercise market power.

Karakatsani and Bunn (2004) also model British electricity prices, and also include the demand/supply ratio in their regression models, along with a wide variety of other variables. Their model for half-hourly prices is structured as a multiple regression, within which they test several variations. For the error process they include both conditional and unconditional processes for heteroskedasticity[21]. They allow the regression parameters to vary over time[22], and they also allow the regression parameters to change depending on the state of the market, with the states following Markov regime-shifting.

The list of variables in their regression models for half-hourly prices includes the following:

- demand
- demand slope (the rate of change in demand)
- demand curvature (the rate of change in the demand slope)
- demand volatility for that period over the past week
- margin (the excess generation capacity: maximum possible output minus demand)

[20] The Lerner Index (Lerner, 1934, as cited in Church and Ware, 2000, p.36) is a commonly-used measure of price mark-up in markets. It is calculated by dividing the difference between price and marginal cost by the marginal cost. The greater the value of the Index, the higher the likelihood that market participants are exercising market power.

[21] They find that the structure of price volatility varies throughout the day.

[22] They conclude that the use of time-varying parameters "seems more appropriate for highly evolving or unstable markets".

- expected imbalance (indicated generation minus predicted demand)
- scarcity[23]
- price volatility for that period over the past week
- the price for the same period the previous day
- the daily average for the previous day

Despite not explicitly specifying a process for jumps in prices, as many other researchers do, Karakatsani and Bunn find that they are able to model prices (and the volatility of prices) extremely well. This is understandable, as they include just about all the factors (except for transmission outages) that could possibly lead to volatility in prices.

Mount, Ning and Cai (2006) present a regime-switching model especially suited to predicting the occurrence of price spikes. They illustrate that the PJM market offer stack is composed of two sections – a flat section for low levels of load, and a steep section for higher levels, with a distinct kink in between. This gives rise to the two regimes in their model, low-price and high-price: "Whenever the forecasted load is higher than the kink in the offer curve, price spikes will occur much more frequently". Instead of using constant transition probabilities between each regime, the probabilities presented (as well as the underlying price level in each regime) depend on both the observed load and the reserve margin, which they calculate as: (total offered capacity divided by load) – 1[24]. They find that the reserve margin is negatively related to the mean prices in both regimes, and the load is positively related to the mean price in both regimes, which are both expected results. Surprisingly, load was not a significant variable in the transition probability equations. But, as expected, "the probability of staying in the low-price (high-price)

[23] To calculate the scarcity, they first calculate the ratio of margin to demand: that ratio is therefore (maximum output minus demand)/demand. The scarcity is calculated as maximum(lower quartile of the ratio minus the ratio, zero). So, if the ratio is smaller than the lower quartile of the ratio, then the scarcity is positive; if not, it is zero. If the scarcity value is greater than zero it represents a higher probability of a jump in prices occurring.

[24] Therefore the greater the difference (or ratio) between the total offered capacity and the load, the greater the reserve margin.

regime is higher (lower) for larger values of the reserve margin". This is in line with their initial observations regarding the composition of the offer stack.

The models in other studies do not include such measures of excess capacity; however, they still include variables representing both sides of the equilibrium equation.

In order to gauge the effect of interaction between three different regions in the north-eastern United States (PJM, New York and New England), Leonard, Reitzes, Schumacher and Bohn (2002) regressed the price from each region on the gas price as well as temperature variables from each of the three regions. They found interaction between the regions only at hotter temperatures – i.e. when demand increased in a particular region, the level of imports into that region would increase the other regions' prices as well as its own price.

Contreras, Espínola, Nogales and Conejo (2003) estimate Autoregressive Integrated Moving Average (ARIMA) models to predict spot prices in Californian and Spanish markets. They added demand as an explanatory variable for both models, and for the Spanish model they also added a variable to represent the daily amount of hydro generation available. They found that including this supply-side information improved the fit of their forecasts in the months in which there was a great deal of hydro generation available, as the spot prices were unusually low in those months.

Gonzalez, San Roque and García-González (2005) use an Input Output Hidden Markov Model to model Spanish electricity prices. These type of models are a variant of neural network models, which combine several series of inputs, and using a layer of modelling hidden from the user, 'train' various parameters representing those input variables to produce a forecast for the price[25]. The explanatory variables used are the load, generation from nuclear, thermal and hydro generators, and several lags of the price itself. They are

[25] Other studies using neural networks to model electricity prices include Ramsay and Wang (1997, as cited in Sansom and Saha, 1999), Sansom and Saha (1999) and Lora, Santos, Santos, Ramos and Expósito (2002).

able to use the model to identify several different states in the time period they analyse, with states changing not when price behaviour changes, but when the relationships between the explanatory variables and the price change. In their final model, the significant variables for each of their four states are the previous period's price, current hydro generation (which has a negative coefficient, as expected), and this and last period's load.

2.4 Comments

A growing number of top-down price models now include exogenous explanatory variables, such as load and generating capacity, that were previously used only as input into bottom-up models of prices. Both demand and supply influence spot price behaviour; therefore it is pertinent to include proxies for both variables in models for spot prices.

Some models include both supply and demand variables, but in such a way that they influence price independently from each other. This will have the effect of modelling changes to the underlying level of prices. However, it is when shocks occur to demand and supply simultaneously that prices are at their most volatile. Other top-down models have incorporated these explanatory variables in ways consistent with the actual operation of electricity markets, identifying that it is the interaction of demand and supply that set the price. Both approaches can produce reasonable results, for the reasons detailed above. However, the approach of Karakatsani and Bunn (2004), in which supply and demand variables are included both independently and in a related manner, seems perhaps the most appropriate method of those employed. In this way, models can estimate not only the underlying level of prices but also the extreme volatility, the critical characteristic of spot price time series to model.

As constrained capacity is an obvious cause of price spikes, including a variable to this effect seems pertinent, and is well-supported in the academic literature. However, the nature of the market producing the spot prices should determine just how the extra information is incorporated. Using a variable that represents system "tightness" through a measure of excess capacity is appropriate for markets dominated by nuclear or thermal capacity, in which the amount of available generating capacity is clearly defined. However even these margins may be misleading, with emissions constraints now limiting the ability for certain types of plant to run, even if they are in the money. In markets that include a high proportion of hydro generation, such as New Zealand's, or when the supply of thermal fuel is limited, other techniques must be employed to measure the availability of generation. As explained further in Chapter 4, this is because while hydro generators can be called on to operate in any given period (provided they have available "fuel"), to do so often compromises their ability to generate in subsequent periods. Calculations of available capacity must take into account future demand and capacity, rather than simply the ability to generate in the current period.

It is clear, therefore, that incorporating extra information into top-down price models in most cases improves their ability to model and forecast spot prices. Just how this information is incorporated, however, should be determined after thorough examination of the actual markets for which prices are being modelled.

3

THE NEW ZEALAND ELECTRICITY MARKET AND ITS SPOT PRICES

3.1 Introduction

While deregulation has been a feature of many electricity markets around the world, the majority of the research presented in this thesis focuses on market behaviour in the New Zealand Electricity Market (NZEM). This chapter follows on from the broad literature review presented in the previous chapter by describing those studies that have examined spot price behaviour in the NZEM. The following section provides background on the NZEM, illustrating spot price behaviour since the formation of the market and explaining the underlying causes of that behaviour. Section 3.3 reviews the studies of NZEM prices that exist in the academic literature, explaining in detail the components of the most prominent models. In the final section of this chapter, we apply one of the models from Section 3.3 to a more recent time series of NZEM prices, and compare and contrast the results with those of the original study. The discrepancy between the two sets of results provides the motivation for the majority of the research presented in the remaining chapters in this thesis.

3.2 Background

In physical terms, the NZEM consists of 244 market nodes, encompassing both the North and South Islands. The network is linked across the two islands by a High Voltage Direct Current (HVDC) transmission line, which stretches from the Benmore node in the middle of the South Island to the Haywards node just north of Wellington. Between the two islands, power is transferred via an undersea cable known as the 'Cook Strait Cable'. The HVDC line links the large scale generating capacity in the South Island to the high load areas of the North Island.

The primary form of electricity generation in the NZEM is hydro (approximately 65% of total average annual output), followed by thermal (30%) and geothermal (5%). There is a small but rapidly growing capacity of wind generation, however the total impact of new wind power investment is limited by the fact that wind output is both variable and unpredictable (Leyland, 2004). Even with such a large proportion of hydro generation, reservoir storage is limited, with a maximum of only around 15% of annual energy able to be held in storage at any one time. Thus the security of supply is highly dependent on having adequate hydro generation capacity, which itself depends on inflows into the reservoirs, which are highly variable. Therefore the security of supply cannot be guaranteed from month to month, let alone year to year. Just one or two months of drought can result in a serious shortage of water in the hydro reservoirs, regardless of the situation beforehand.

The reformed wholesale electricity market of New Zealand began trading in October 1996, but has since undergone significant changes in terms of structure. Contact Energy, which began operations in April 1996 in order to compete with the Government's Electricity Corporation of New Zealand (ECNZ), was privatised in April 1998, and in April 1999 ECNZ was split into the three state-owned enterprises (SOEs) to compete with Contact Energy. Also in 1999, new legislation came to pass allowing every consumer in the country to choose from which retailer they purchased their electricity.

Like other reformed electricity markets around the world, the NZEM market has not yet reached maturity. In almost every year since 1999, the NZEM has seen significant market events and changes, each of which has had a major influence on wholesale electricity prices. After the 1999 split of ECNZ there was a fight for market share, with each firm bidding as low as possible to try to generate as much electricity as they could. Also that year, legislation was passed allowing generating companies to acquire retail firms, which, as some of them did so, could explain changes to the competitive bidding behaviour. At the start of 2000, the Otahuhu B generating plant came online, easing transmission congestion in the upper North Island and thereby reducing spot prices in that part of the country. Also that year, industry sources suggest that the generating firms began to gain more of an understanding of how to take into account their contract levels when bidding. The increased price volatility reflected the shape of the offer stacks developed during this period.

In winter 2001, for the first time since 1992 a serious drought resulted in a shortage of reservoir storage, which caused a significant and sustained increase in spot prices. However in mid-2001, the distribution of retail contracts between the market players changed markedly; one retailer was forced to withdraw from the market after it had not hedged against the high prices, with an existing generating company acquiring all that retailer's contracted customers. Perhaps coincidentally, spot prices fell significantly almost at the same time, without there being any change in the storage situation. The prices in 2002 were relatively low and stable as inflow levels were high; however in 2003 another water shortage, combined with concern over an adequate coal supply (Daniels, 2005) caused capacity to be severely stretched, and prices were again high. In response to the 2003 water shortage, the Government established the Electricity Commission to oversee the operation of the NZEM (among other roles), though the Government stated they would not regulate prices (Steeman, 2004). Instead, they constructed a high marginal cost gas thermal generation plant in Whirinaki, to be run to relieve extreme hydro storage situations and other market conditions. In late 2005 and early 2006 another water shortage (see Gorman, 2005) required that plant to be dispatched (see Gorman and Steere, 2005), as spot prices again exceeded \$200/MWh for a number of weeks[1].

The behaviour of NZEM spot prices since deregulation, including the high-price events of 2001 and 2003, can be observed clearly in Figure 3.1 below, which shows the average NZEM spot price per day from 1 October 1996 to 30 June 2003. Most of the research into NZEM spot prices presented in this thesis involves data from the break-up of ECNZ in 1999 until the end of June 2003, hence the series of spot prices in Figure 3.1 is truncated at this point.

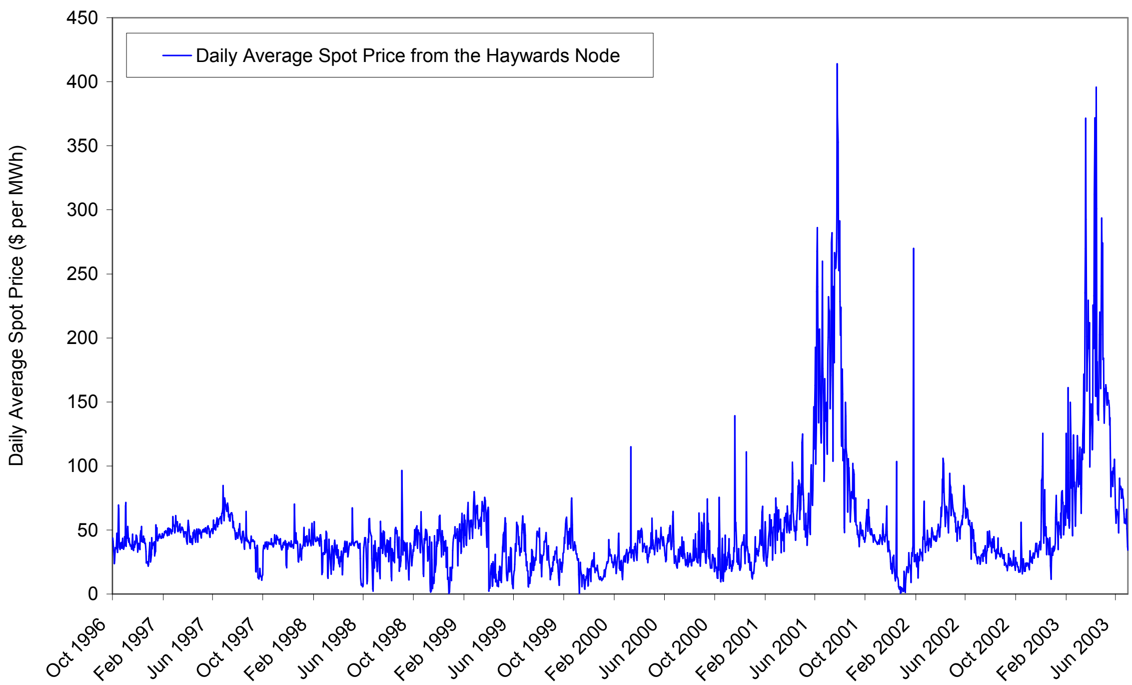

Figure 3.1: New Zealand daily average spot prices at the Haywards node, 1 October 1996 – 30 June 2003

In the period between the formation of Contact Energy in 1996 and the formation of the three SOEs in 1999, spot prices were relatively stable and seldom high, due largely to the

[1] Several of these market situations are examined in depth in this thesis. However, for more information regarding the history of the NZEM and its price behaviour, see Mason (2002) and Videbeck (2004).

fact that there were no periods of sustained low inflows. Unlike markets dominated by thermal generation, in which the spot price may be extremely volatile in the short term but less so in the long term, the NZEM spot prices can reasonably be expected not to fluctuate greatly from day to day. This is especially evident in the first two years of the market, where the price has only small variation about a level. Price movements can be long-lasting, as can be seen in both 2001 and 2003, which suggests that changes in supply- and/or demand-side factors have lasting effects on the price in the NZEM[2].

3.3 Previous research into the behaviour of NZEM spot prices

As indicated above, due to the dependence the NZEM has on hydro generation with limited seasonal storage, New Zealand's spot prices exhibit characteristics quite unlike prices from thermal markets. The majority of price models discussed in the previous chapter therefore are not transferable to the NZEM, as they cannot account for the periods of sustained high prices induced by water shortages. Several studies have been undertaken into modelling New Zealand's spot prices; however, to our knowledge, none of these studies extend to modelling the unusual price behaviour in the two "dry" years of 2001 and 2003 by way of incorporating information on the reservoir storage level. Several of these previous studies are detailed in this section.

3.3.1 Mason (2002)

In his 2002 Masters thesis, Mason analyses prices from both the Benmore and Haywards nodes between October 1996 and June 2000, in both high (half-hourly) and low (daily average) frequency. Initially, he examines the intra-day patterns of prices at these two nodes, and finds that, in general, volatility is greater at the Haywards node than at Benmore. He then proposes individual models for each half-hourly price on any given day. He uses an eight term vector autoregression (VAR) model to construct the 48 price models, with his key finding being that the explanatory power of these models is less

[2] It should be noted that these effects are of a seasonal nature, rather than affecting the spot price directly in the long-term.

during the peak periods of the day (periods 15-17 and 36-41). Therefore, previous prices offer less insight into future prices in peak periods compared with off-peak periods.

After discussing the literature on spot price modelling available at the time of his study, Mason presents a two-factor model of daily average NZEM spot prices. The first factor is the process followed by a central tendency, which is slowly mean-reverting with low volatility. This factor accounts for the underlying mean level of prices, which shows gradual variation over time. The second factor accounts for the fluctuation of prices around and away from the central tendency, and is characterised by high volatility but rapid mean-reversion. He tests the hypothesis that the parameters of these two processes vary across several pre-defined sub-periods in the sample period, and perhaps the most important result he found was that while the central tendency process appeared constant throughout the sample period, the estimated parameters of the short-term volatility process varied significantly over time.

Mason's results and conclusions are consistent with those of other studies. Regarding the central tendency, he states that, "In the long-run, [it] reverts to a static value that is principally determined by costs of production"[3]. With this in mind, it is unfortunate that his research was undertaken before 2001, when the short-run costs of production rose dramatically. The second factor in his model represents the short-term volatility inherent in spot price time series; however it does not make an allowance for extreme volatility, such as price spikes.

[3] This is a common view for the reason why prices revert to a specific value over time. The value depends on the long-run marginal cost (LRMC) of production, rather than the short-run marginal cost (SRMC). This is because in the long term, generation will be offered at price levels that recoup not only the actual cost of generation (the SRMC) but also the cost of constructing generating plant, both of which are included in the calculation of the LRMC of generation.

3.3.2 Guthrie and Videbeck (various)

Guthrie and Videbeck (2002b) and Videbeck (2004) propose models for half-hourly spot prices, as opposed to daily average spot prices, in order that their models may be of use, for example, in short-term decision-making and pricing specific financial derivatives. As they note, high-frequency price models are somewhat less prevalent in the academic literature, despite their relevance and function. Motivated by the fact that electricity cannot be traded across time, they propose that electricity generated at different times of the day can be treated as different commodities.

The authors examine prices from the Haywards node in two separate periods: 1 March 2000 – 28 February 2001, and 1 March 2001 to 28 February 2002. Their sample period includes the "dry year" of 2001, and they note that their results for that year were influenced by the extreme hydrological conditions. As a result, much of their attention focuses on the results of their first year's price data, as they believe 2001 was "not representative of a regular year". While they do include some analysis of the effects of the drought on their results, the models they develop do not take into account the reasons behind the high prices. Perhaps if they were to have undertaken their study a year later and included price data from 2003 (another non-"regular" year), their model would have incorporated a measure for the changes in the underlying price level and price dynamics.

Their 2002 paper begins by examining the intra-day correlation structure between the prices in each of the 48 half-hourly periods in their two sample years. They identify that the overall market data may be segmented into four intra-day markets, each relating to a set of trading periods such as "peak", "off-peak", et cetera, based on that structure. On any given day, the prices within each of these markets are highly correlated, however the correlations between individual periods' prices in other markets on the same day is low. This result is both interesting and crucial, as it later allows them to model four prices per day instead of 48, resulting in a significant increase in parameter estimation efficiency.

In an approach similar to that of Mason (2002), their initial model is a periodic autoregression (PAR) model, with each half-hourly price being a function of the 48 half-hourly prices in the preceding 24 hours, as well as daily and monthly dummy variables to account for potential longer-term seasonality in the prices. They then utilise the intra-day market structure to reduce the number of price equations from 48 to four, and extend their model so that "the base price in any market depends on the base prices in the previous four intra-day markets[4] and a noise term", with the noise term variance varying by intra-day market. The price for an individual half-hourly period therefore becomes the sum of:

- a constant price component for that half-hourly period (the same across the whole year)
- constant price components for that day and for that month (to account for the seasonality)
- the base price component for that intra-day market
- a noise term, the variance of which varies depending on the half-hourly period

The estimated parameters of this state space model reveal interesting characteristics of the influence of different markets' prices on each other, such as the fact that the morning peak price level depends on both the preceding overnight price and the previous evening's peak price[5]. Not surprisingly, they find that prices in the peak periods are higher and more volatile than in other periods. They also observe that, in their sample data, prices are lower in the months from September to November, which is evident in Figure 3.1.

At the end of 2004, now having six years of market data to analyze, Guthrie and Videbeck note in a third paper that "the half-hourly trading periods fall naturally into five

[4] This refers to the four previous intra-day markets in the past 24 hours, rather than the same intra-day market over the past four days.

[5] The overall result is interesting, as it suggests that although individual prices in an intra-day market are unrelated to individual prices in other intra-day markets, the average level of prices in an intra-day market is in fact related to the average level of prices in other intra-day markets.

groups corresponding to the overnight off-peak, the morning peak, daytime off-peak, evening peak, and evening off-peak”. Instead of using raw price data, as in 2002, the authors this time analyze the residuals from a filtering regression to remove weekly and monthly seasonal patterns. They present the same PAR models as in their 2002 paper; however, they now draw conclusions regarding the timing and extent of volatility in prices throughout the day. They find that price shocks (e.g. spikes) are larger but less persistent in peak periods compared with off-peak periods, but that the peak shocks may reappear in subsequent peak periods.

Overall, Guthrie and Videbeck offer sound analysis on the intra-day structure of NZEM prices. Their high-frequency models offer a useful extension to the models for daily average prices presented in the following chapters of this thesis.

3.3.3 Escribano, Peña and Villaplana (2002)

In their 2002 paper, Escribano, Peña and Villaplana (EPV) compare the performance of their set of models when applied to time series of daily average prices from several markets around the world, including prices from the NZEM’s Haywards node from 1 October 1996 to 31 August 2000. As shown in Figure 3.1, prices during this period are very different to prices from August 2000 onwards.

The approach taken by EPV follows a similar reasoning to that of Mason (2002) and other authors, including Lucia and Schwartz (2002). EPV use the basic hypothesis that a spot price time series (P_t) can be decomposed as the sum of two distinct components – a deterministic component ($f(t)$) and a stochastic component (X_t), as shown in equation [1]. Through these two components, they aim to account for each of the features inherent in electricity spot price time series. The deterministic component models the underlying or expected price level for any given day, accounting for price seasonality, and the stochastic component models the complex volatility structure of electricity prices.

$$P_t = f(t) + X_t \qquad [1]$$

In their model, the deterministic level of the electricity price series $f(t)$ is estimated essentially using a linear step-function regression[6], made up of a constant, trend, and seasonal variables, as shown in [2]. The regression includes binary dummy variables which represent the month of the year, and whether or not the particular day is a weekday (i.e. Monday – Friday), or a weekend. These dummy variables take into account the seasonal patterns in the price level due to the time of year and the day of the week. The constant represents some base level of the price, and the linear trend aims to capture any long-term price movements that may be present.

$$f(t) = Const + Trend.t + \sum_{i=2}^{12} M_i D_{i,t}^{M} + wkd.D_t^{W} \qquad [2]$$

where: *Const*, *Trend*, the M_i and *wkd* are estimated coefficients

D^M and D^W are Monthly and Weekday dummy variables, respectively

Added to the deterministic price level is the stochastic component X_t, which aims to model the particular volatility structure of electricity prices. The stochastic component can itself be decomposed, this time into the sum of three separate contributors to volatility – autoregression (AR), "usual" volatility and "jumps", as shown in [3].

$$X_t = \theta X_{t-1} + \text{"}usual\text{"} + \text{"}jumps\text{"} \qquad [3]$$

The AR component is included because of the persistence in the effects of shocks to electricity price series. Generally, the effects of a shock to the price series, large or small, will last for more than one subsequent period, and θ is usually positive. The effect of any particular shock is dampened over time, which leads to the price level reverting back to

[6] In other versions of their working paper, they also model the deterministic component using a sinusoidal function. They state that the sinusoidal function is more appropriate for modelling price series which contain a regular seasonal pattern, which, from Figure 3.1, is evidently not the case for the NZEM.

its base (or average) level, as discussed in Chapter 2. The size of θ determines the speed at which the price reverts back to its mean level after a shock.

In their paper, EPV model "usual" volatility as a General Autoregressive Conditional Heteroskedasticity (GARCH) process. This type of process is used for modelling non-constant variability in the residuals of a model, known as heteroskedasticity. In electricity price time series, there are both periods of prolonged high variance and periods of prolonged low variance. A GARCH process also takes into account persistence in the volatility through its AR component.

A GARCH(1,1) process suggests that the "usual" volatility, e_t, is a "real-valued discrete-time stochastic process" (Bollerslev, 1986). Each realisation of the e_t process is conditional on every previous realisation, and is distributed normally with a mean of zero and a variance h_t, with h_t modelled by:

$$h_t = \omega + \alpha\, e_{t-1}^2 + \beta h_{t-1} \qquad [4]$$

Equation [4] states that this period's forecasted variance in the residuals is a function of an underlying average (or unconditional) variance, ω, last period's forecasted conditional variance (h_{t-1}), and the square of last period's actual movement from the expected value (e_{t-1}^2). The ARCH term (α) is included to account for short-term clustering and persistence in the volatility of the residuals, whereas the GARCH term (β) contributes more to long-term volatility persistence[7]. The GARCH process is particularly relevant for modelling electricity prices, due to the fact that they exhibit patterns in both the short-term (due, for example, to temporary transmission outages) and the longer-term volatility (due to plant maintenance or seasonal shortages in supply). Estimates are required for ω, α and β in the GARCH(1,1) process.

[7] See Engle (2001) for further explanation of the GARCH parameters.

The EPV model incorporates jumps in the price series through a Poisson jump process. Each day, there is a certain probability λ_j of a jump occurring. If a jump does occur, then its size is drawn randomly from a normal distribution with mean μ_j and standard deviation σ_j. The jump probability and the jump size mean and standard deviation may all vary depending on the season *j*, and each of these parameters requires estimation. If no jump occurs, then the estimated volatility is just the sum of the autoregressive component and the "usual" volatility.

In summary, the stochastic component of the price model can be represented as follows:

$$X_t = \begin{cases} \theta X_{t-1} + e_t & \textit{with probability } (1-\lambda_j) \quad \textit{No Jump} \\ \theta X_{t-1} + e_t + \mu_j + \sigma_j \varepsilon_{2t} & \textit{with probability } \lambda_j \quad \textit{Jump} \end{cases}$$

where $e_t = \sqrt{h_t}\,\varepsilon_{1t}$, and $\varepsilon_{1t}, \varepsilon_{2t} \sim \mathrm{N}(0,1)$. [5]

This model is very detailed and requires the estimation of many different parameters. The parameters are estimated concurrently using a technique called Constrained Maximum Likelihood estimation (CMLE). This technique, which is detailed more extensively in Appendix B, involves finding the set of parameters for the model that is *most likely* to have produced the actual series of data observed. For the estimation of this model, the CMLE procedure is constrained by the fact that all variances must be positive, and all probabilities must be between zero and one. There are no other constraints or assumptions on the distributions of any of the other parameters.

EPV estimate a set of nested models consisting of the deterministic component and various combinations of the stochastic parameters. The models they most prefer are the more complex ones, containing all the components listed in [5], as opposed to the models without the time-varying jump parameters or the GARCH specification for the residual error.

They find that, unexpectedly, the price series generated by the New Zealand and Scandinavian markets, while both dominated by hydro production, behave quite differently. Prices from Nordpool have a lower degree of mean-reversion (i.e. a larger θ), which is expected due to the fact that hydro reservoirs act as indirect storage of electricity and therefore allow a degree of inter-temporal substitution to smooth out shocks in demand and supply. For example, in the medium-term water can be conserved in periods of low demand and used when demand is higher, which effectively maintains the amount of excess generating capacity available at a constant level over time. In the short-term, surges in demand or sudden supply outages do not have the major effect on prices that they do in thermal systems, due to the relative flexibility of hydro generating units which usually do not require long periods to ramp up.

However, shocks in hydro supply due to extreme weather conditions can have lasting price effects in the medium-term, as a reduction in generating capacity induced by low reservoir levels can take weeks or months to correct. Therefore while short-term price volatility may be lower in markets with a high proportion of hydro power, average spot prices are less stable (a feature noted also by Wolak, 1997; Aires, Pereira, Lima, Barroso & Lino, 2002; Audet, Heiskanen, Keppo & Vehviläinen, 2004; and Lucia and Schwartz, 2002). EPV note this in the case of Nordpool, for which a dry year in 1996 resulted in higher average prices, but the NZEM prices in their sample did not appear to be affected by such extreme weather patterns, and thus had a high degree of mean-reversion[8].

3.3.4 Other studies

Several other studies and analyses of New Zealand's electricity spot prices exist, however they analyse and make observations regarding the price behaviour rather than attempting to model this behaviour. Some of these studies are described in this section.

Wolak (1997) examines the behaviour of prices in several electricity markets, and relates their behaviour to the structure of the individual markets. He fits VAR models to prices

[8] It should be noted, however, that while hydro storage in both Nordpool and the NZEM is highly seasonal, Nordpool's reservoirs are much larger than those of the NZEM.

from all of the markets he studied, except for New Zealand's, due to the fact that, at the time of his study, he only had access to less than a year's worth of data from 1996.

In their 2004 paper, Wilson and Cheng ask the question "Is the electrical spot price chaotic?" Their analysis reveals that the time series of NZEM spot prices from the Benmore node does, at times, exhibit chaos. The implications for efforts to model and forecast the spot price are not entirely clear from the paper, and the authors do not propose any particular models themselves for the spot price series. They do suggest, however, that while the usefulness of studying chaotic systems is not questionable, any forecasts from such systems may be.

Li and Flynn (2004a, 2004b) make observations regarding the intra-day (diurnal) pattern and volatility of New Zealand's spot prices at the Benmore node, as well as prices from thirteen other markets. Their sample included half-hourly NZEM prices from November 1996 to December 2001. Their findings include that:

1. Price volatility is higher on weekdays than on weekends in all fourteen markets.
2. Intra-day price volatility in New Zealand is lower than average when compared with the volatility in other markets.
3. On average, New Zealand prices have two peaks per day, the timing of which, unlike other markets, does not vary depending on the day of the week.
4. In most of the markets they studied, the correlation between prices and load is low; in only three of the thirteen markets is the correlation greater than 0.4. However the authors were unable to source a series of load from the NZEM, and were therefore not able to compute the correlation for this market. Overall, this is a somewhat surprising and important result, and suggests that load may not be an important explanatory variable in high-frequency models for spot prices. However, empirical evidence suggests that when NZEM load increases prices do rise as a result, especially when the supply-side is stretched ("Power prices rise amid cold snap", 2004).

At the time we commenced our research, the EPV model was one of the most sophisticated top-down models available for modelling daily average spot prices, and it still is. The paper introducing their model is one of the most widely-referenced in the spot price modelling literature. As our research involves incorporating features from a bottom-up model into a top-down model for daily average spot prices, rather than developing a new top-down model, we required an existing model to use as our base. The EPV model accounts for all features commonly recognised in electricity price time series, including long-term trends, seasonality, heteroskedasticity and jumps. For these reasons, we chose the EPV model as our base, and began by fitting it to a more recent series of prices than the developers had in their paper.

3.4 Modelling the NZEM spot prices from 1999-2003

The time series EPV use ended prior to 2001, and encompassed a period in which the deregulated market was still in its infancy and prices were relatively stable in New Zealand (see Figure 3.1). Using the CMLE routine in GAUSS 6.0, we fit the EPV model detailed in Section 3.3.3 to a different price series, daily average spot prices from the Haywards node from 1 August 1999 to 30 June 2003. Figure 3.2 and Table 3.1 below, of the respective empirical distributions and tables of summary statistics of the two time series, illustrate the differences between these series. The graph on the left (1996-2000) has a near symmetric distribution and is almost mesokurtic, as indicated by its skewness and kurtosis both being close to zero. In contrast, the 1999-2003 price distribution is positively skewed and has relatively large tails (leptokurtic). Both have a similar median price level, however the average price of the later time series is much greater due to having a fat tail extending toward higher values.

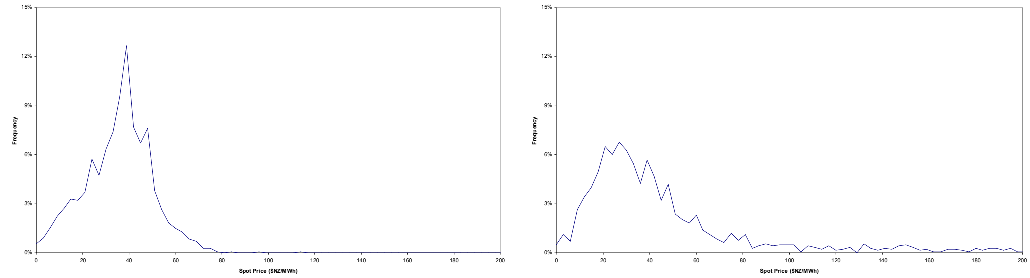

Figure 3.2: Empirical distributions of daily average spot prices from the Haywards node: October 1996 – August 2000 (left) and August 1999 – June 2003 (right)

Series	Obs	Mean	Median	Min	Max	Std Dev	Skew	Kurt
1996-2000	1431	37.05	38.20	0.58	115.13	14.31	0.05	0.63
1999-2003	1430	56.51	39.88	0.61	413.97	53.82	2.82	9.55

Table 3.1: Summary statistics for daily average spot prices from the Haywards node: October 1996 – August 2000 and August 1999 – June 2003

EPV present an exhaustive list of appropriate tests to detect the presence of a unit root in a time series, however in this study we limited ourselves to the Augmented Dickey Fuller test for stationarity[9]. In their study they found conclusive statistical evidence that the prices in the NZEM were stationary, however in our time series a p-value of 4.8% was almost not enough to reject at the 5% level of significance the hypothesis of the presence of a unit root in the data. Having a time series generated by a unit root process is an undesirable property, as it essentially implies that the series is theoretically unbounded (see Robinson (2000) for further details) and will not necessarily revert to an underlying level.

[9] The Dickey Fuller test, developed by Dickey and Fuller (1979, as cited in Makridakis, Wheelwright & Hyndman, 1998, p.329) tests the null hypothesis of the presence of a unit root in a time series. If a unit root exists in a time series, the data must be differenced first before being stationary. Stationarity is usually a requirement for unbiased estimation of most time series models. The Augmented Dickey Fuller test, developed by Said and Dickey (1984), should be used instead if the data is linearly trended.

As shown in Table 3.2 below, the estimated coefficients from our sample are extremely different to those estimated in the EPV study. The deterministic component is at a much higher level in the EPV sample, and the greater significance (in terms of t-values) of the monthly dummy variable coefficients suggests that the seasonal pattern was much more regular in the EPV sample than the later sample. The estimated trend coefficient is negative in the EPV sample, which would have come as a result of the prices from June 1999 – August 2000 being lower on average than in the first half of the sample (recall it has been suggested that this period was characterized by competitive bidding). The weekday/weekend effect was similar in both samples; on average the prices during the week are around $3 greater than prices on the weekend.

	from Escribano et al. 1 Oct 1996 - 30 Aug 2000			Updated study 1 Aug 1999 - 30 June 2003	
	Coefficient	t-stat		Coefficient	t-stat
CONSTANT	44.07	54.09		18.67	6.17
TREND	-0.016	-30.23		0.011	3.33
M_2	4.01	4.78		4.41	1.73
M_3	6.20	6.39		10.83	3.64
M_4	2.57	2.62		16.97	5.10
M_5	5.88	6.40		13.24	3.20
M_6	12.03	11.80		7.47	1.96
M_7	5.53	4.90		6.35	1.84
M_8	2.26	2.04		-0.10	-0.03
M_9	-2.90	-2.58		-0.44	-0.14
M_{10}	-0.21	-0.21		2.30	0.71
M_{11}	-1.32	-1.24		0.66	0.20
M_{12}	-5.31	-5.81		-1.25	-0.64
WEEKDAY	3.00	10.80		3.29	10.32
λ_{AUTUMN}	9%	3.11		6.78%	3.18
λ_{WINTER}	-2%	-0.45		4.96%	2.57
λ_{SPRING}	-2%	-0.65		2.54%	2.31
λ_{SUMMER}	1%	0.17		5.87%	3.21
θ	0.57	22.21		0.89	75.55
ω	1.17	3.61		5.05	5.59
α	0.38	8.74		0.38	8.45
β	0.58	19.39		0.51	16.50
μ	3.59	1.45		31.59	4.39
σ	16.31	18.05		33.01	4.51
Log-Likelihood	-4948			-5113	
SIC	10077			10401	

Table 3.2: Estimated parameters from application of EPV model to daily average spot prices from the Haywards node: October 1996 – August 2000 (from EPV paper) and August 1999 – June 2003

The estimate of the daily constant unconditional variance parameter, ω, is much larger in the later sample than in the EPV sample. This suggests not only that prices are more volatile now compared with the early days of deregulation, but that the low storage levels could have caused an increase in not only the overall price level but also the “usual” daily volatility. The estimated jump coefficients are significant in each season in the later sample, with both the jump mean and standard deviation being significantly larger. These results are consistent with EPV, as they find that price series whose distributions have a larger degree of skewness have higher jump means.

Possibly the most interesting dissimilarity in the results of the two samples comes in the comparison between the two autoregression parameters (θ). This brings the results from New Zealand much closer into line with those from Nordpool in the EPV study, in that prices take much longer to revert back to their underlying levels in the second sample than they do in the EPV sample. This is the result expected (but not realised) by EPV in their study, due to the NZEM being dominated by hydro generation[10].

An alternative interpretation of this result is that the deterministic price level from 1999-2003 is not well modelled by the estimated constant, trend and dummy variables, as a less accurate deterministic price path would be evidenced by a mean-reversion parameter closer to 1. If instead the deterministic price path fit the actual price path almost perfectly, the mean-reversion parameter would be close to zero, as stochastic shocks to the price would create a deviation away from the deterministic price path for the current period only. Figure 3.3 below shows the time series from the second sample with the estimated deterministic component overlaid.

[10] It could also be the case that from 1996-2000 there were no events that caused prices to deviate markedly from their long-run average levels.

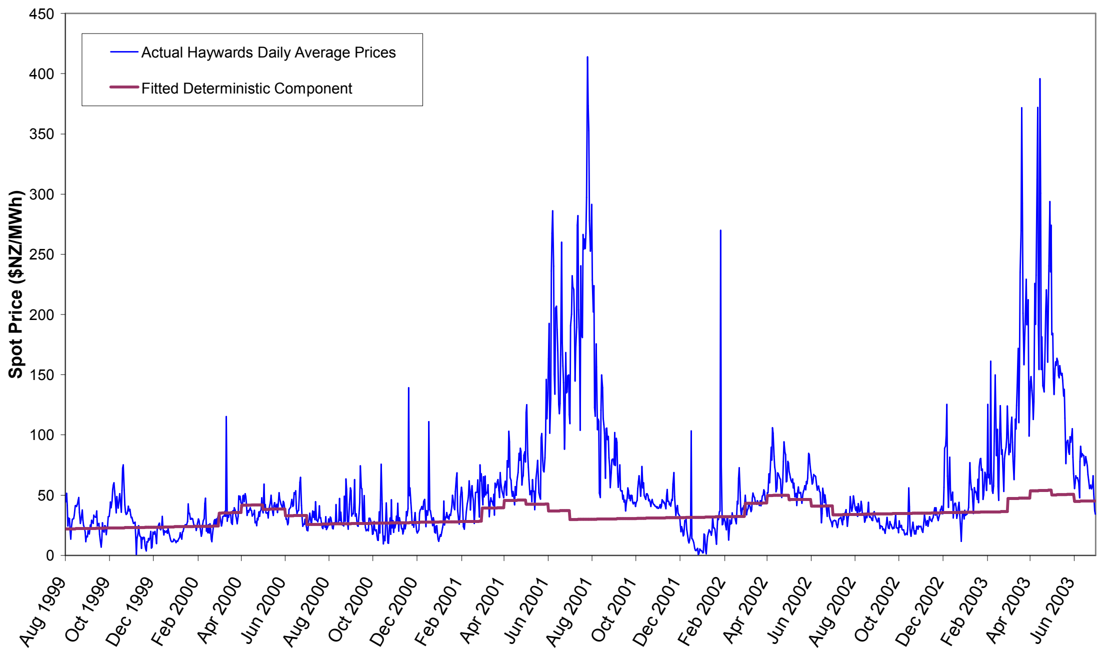

Figure 3.3: Daily average spot prices from the Haywards node, August 1999 – June 2003 and estimated deterministic component (less weekend effect) from the fitted EPV model

The monthly dummy variables are intended to capture the seasonal patterns in the time series, and indeed appear to do so in 2000 and 2002, which could be classed as "regular" years. However, the fit to the underlying price paths of 2001 and 2003 (when the storage level was low) is not good; all price behaviour above the stepped line must then be accounted for in the stochastic component. While it is unlikely to be avoidable with this particular data set, there will obviously be major inter-temporal differences in the residuals used to estimate the coefficients of the stochastic parameters. It would therefore appear necessary to improve the model for the deterministic component by including information that reflects the relationship between the storage level and the underlying level of the spot price.

As no other studies in the academic literature to our knowledge have attempted to model a time series of NZEM prices by incorporating a measure of the storage level, a new approach is required. Having illustrated the inability of a complex top-down price model to mimic NZEM price behaviour since 1999, we now show in the following chapter how the fit to prices can be improved by incorporating storage level information.

4

THE INCORPORATION OF HYDRO STORAGE INTO A MODEL FOR NEW ZEALAND SPOT PRICES

4.1 Introduction

As was illustrated in the previous chapter, New Zealand's spot prices appear to have been, and typically will be, heavily influenced by the level of hydro storage available for generation. This is due to the heavy reliance the country has on hydro power as a source of generation, and the fact that storage for that generation is limited. Because of this, even a month of low inflows can result in an increase in spot prices, and, vice versa, a week of rain during one of these periods can dramatically ease an otherwise tight storage situation and reduce prices (see "Heavy rains help ease power prices", 2006).

Recent top-down models of spot market prices aim to account for all the characteristics exhibited by electricity price time series: seasonality, mean reversion, price jumps/spikes and time- and price-dependent volatility. They are particularly well-suited for thermal-dominated systems, in which extreme short-term price volatility and strong mean-reversion are dominant characteristics. However, as illustrated in the previous chapter, even these models can be shown to fit poorly to price behaviour in New Zealand's market. This is due to the large proportion of hydro generation (around 65% of average annual output) and relatively small long-term reservoir storage capacity (around 15% of average annual output).

As explained earlier, in the nine years since the formation of a wholesale electricity market in New Zealand in 1996, there were two "dry" years, 2001 and 2003, in which limited flows into the hydro reservoirs resulted in episodes of sustained high spot prices. While the reservoir levels were clearly low at these times, public debate still exists as to whether prices were pushed high through manipulation or whether the high prices were simply accurate signals to the market that there was a serious shortage of supply.

In order for a top-down price model to perform well in the New Zealand market, and be useful for both regulation and strategy-planning during these dry years, it needs to be modified to incorporate the physical factors that have such a powerful influence on the price level. Hydrological factors, such as storage levels and inflows, are major drivers of hydro generator behaviour. Assuming generators use modern reservoir management optimisation theory, both factors are taken into account in the calculation of marginal water values (MWVs), which, theoretically, form the basis of their supply offers (see Yang, 1995, Scott 1998, Baíllo, Ventosa, Ramos, Rivier and Canseco, 2001 and Johnsen, 2001, among others). Once the MWV has been calculated, the hydro generator can be operated in the same way as a thermal generator, treating water as its "fuel", and the MWV is the price at which it must purchase water from the reservoir. Only when the marginal benefit of releasing water exceeds the MWV will the generator be dispatched. The MWVs are assessed internally however, and are not public knowledge; therefore in order to model spot prices using the storage level some proxy for the MWV is required.

In this chapter, we use reservoir management theory to extend the EPV model that was shown in Chapter 3 to be unable to account for the extreme price behaviour experienced in 2001 and 2003. The storage level is transformed into a crude measure of the MWV, and incorporated into the model as the major driver of the deterministic price level. Our analysis shows that spot price levels in New Zealand (especially during dry years) can be modelled with increased accuracy using the storage level, and this gives some explanation for the periods when prices are sustained at levels higher than their long-run average. These results are of significant value for risk management and regulation in not only the New Zealand market, but in any market in which a large proportion of supply is met through hydro generation.

4.2 Modelling prices influenced by hydrological factors

For many electricity markets around the world, hydro generation makes up a significant proportion of total generation. These markets will no doubt experience the same problems as the NZEM, in that reductions in storage reservoir inflows will decrease offered hydro generation and likely increase the spot price. However, in most other hydro-dominated markets, storage is not as limited as it is in the NZEM. In recent years, several econometric models of prices from such markets have found that including explanatory variables related to hydro generation has improved the performance of the models. However, as explained in the previous chapter, to our knowledge no such study exists for the NZEM.

In the Nordic region, over 90% of generation comes from hydropower (Gjolberg and Johnsen, 2001). As mentioned in Chapter 2, several studies exist in which storage levels were incorporated as supply-side factors in models for the Nordpool spot price. Botterud, Bhattacharyya and Ilic (2002), Gjolberg and Johnsen (2001), Johnsen (2001) and Koopman, Ooms and Carnero (2005) all make mention of the influence of storage levels on prices in the Norwegian electricity market, and the latter three studies each find some measure of storage (or inflows) to be a statistically significant variable in their price models.

Gjolberg and Johnsen (2001) model the difference between forward and spot prices as a function of the storage level (measured as a percentage of total storage capacity), in monthly intervals. They state that forward prices are based largely upon long-term expectations of storage levels. As a result, they find that when storage levels are emptier (more full) than expected levels, spot prices are higher (lower) and the difference between forward and spot prices is lower, sometimes negative (greater, positive). Instead of using storage levels, Johnsen (2001) models the change in price from week to week as a function of the change in the level of inflows. If inflows stay constant from week to week then prices do not change. If inflows increase then the price decreases, but if inflows are high in one week and decrease in the next, then the price increases. Interestingly, he finds that changes in the price are more accentuated later in winter, when storage levels are lower, than at the start of winter.

In a similar study to Gjolberg and Johnsen, Koopman et al. (2005) find that an increase in the weekly reservoir level (again, measured as a percentage of total storage capacity) has "a significant negative effect on electricity prices, except on Mondays, when the measurements for the new week are not yet publicly available". They estimate a model for prices without any storage level information, in which the seasonal element in the spot price time series is modelled with a deterministic time-based component, similar to that of Escribano et al. (2002). However, when they include their measure of the reservoir storage level, they find that the time-based deterministic component for the price series is largely replaced by the annual seasonal cycle of the reservoir level[1].

As already mentioned in Chapter 2, in their 2003 application of ARIMA models to electricity market price time series, Contreras et al. include both load and "available daily production of hydro units" in their models for the Spanish market. They find that their explanatory variables are "only needed in the months with high correlation between

[1] Koopman et al. also include electricity consumption as an independent explanatory variable, and find that it has a significant, positive relationship with the spot price, as expected.

available hydro production and price”. This was because during the months where the availability of hydro increased (possibly through high rainfall, suggesting the ensuing flows could not be stored) the Spanish price plummeted – the opposite case to 2001 and 2003 in the NZEM.

4.3 The Marginal Water Value

In a wholesale electricity spot market, the marginal costs of electricity influence the prices at which generating capacity is offered into the market. For a thermal plant, the costs of generation are relatively clear-cut – the input fuel is purchased and the cost of combustion can be estimated. For a hydro plant, the input “fuel” is water, and in order for the efficient dispatch of generation some assessment is necessary of the (marginal) value of water in storage that the generator can “buy” from the reservoirs as fuel for generation. The concept of valuing water for hydro generation is not recent, and can be traced back to Stage and Larsson (1961, as cited in Read, 1984, p. 6) and Adnét, Auges and Dupoux (1968, as cited in Read, 1984, p. 6), among others (see Laufer and Morel-Seytoux (1979), and Read (1979, 1984)). More recent research by Scott and Read (1996), Scott (1998), and Batstone (2003) has shown that the concept applies in a market context, where participants are expected to “game” their market offers, and must account for the potential gains from doing so, both present and future, when determining the marginal value of the water held in storage.

4.3.1 Definition

The idea of a marginal water value (MWV) is conceptually simple. In any period, the aim of the hydro reservoir manager is to maximise the sum of the profit from release during the period and the value of water remaining at the end of the period. The manager is therefore faced essentially with one basic problem: how much water to release during the period, or conversely, how much water to store up for release in the following periods. As Scott (1998) notes, the MWV is “the price at which water will be traded between the two problems”, and the manager will choose different amounts to release/store depending on the MWV. In a deterministic optimisation, this MWV appears as the shadow price on the water conservation equation.

Alternatively, Guthrie and Videbeck (2002a) and Videbeck (2004) give a description of the marginal cost of hydro generation in terms of real options analysis. On any given day, the hydro generator holds a portfolio of real options, such as the option to generate today, the option not to generate today, and the options to generate or not on any day in the future. As storage is limited and inflows stochastic, generating today can compromise the ability to generate in the future. Therefore, the marginal cost of hydro generation includes not only the physical cost of passing water through the turbines, but also the value of options both created and destroyed by generating today[2].

Valuing water for hydro generation may be complicated further if the water held in the reservoirs is valuable to other users, for example farmers needing the water for irrigation. In situations where this is the case, a generating company will also have the options of selling water now or in the future to other users, rather than have passing it through the turbines and out of the reservoir. The value of these options will have to be included in their calculations for the MWV as well. The largest hydro generating company in the country, Meridian Energy Limited, recently held "informal discussions" with the Government over the possibility of adopting the system of tradable water rights used in Australia (MacDonald, 2004). But water in New Zealand currently may not be traded between users, thus any requirements for other users, along with environmental flow/storage limits, are expressed as constraints, and implicitly priced by the constrained optimisation which produces the MWV.

[2] It should also be noted that only in a perfectly competitive market will a firm's MWV equate to the marginal thermal production costs of that firm or those of competing firms. In a market in which there is some degree of market power, the MWV will reflect the marginal revenue of the hydro firm, rather than its marginal production costs. In such situations, the MWV may change as a result of a change in the structure of the market, even though there may have been no change to market demand or to the supply mix.

4.3.2 Calculation

The optimal storage path in the absence of uncertainty is the one that holds the MWV constant, unless storage bounds make this impossible[3]. And, if storage bounds are reached, they imply high MWV in high demand/low inflow periods (e.g. winter in New Zealand), during which storage is run down from maximum to minimum levels, and conversely in low demand/high inflow periods. This provides a point estimate of the MWV for both the assumed starting storage, and storage along the optimal deterministic trajectory.

In the real world, though, uncertainty is pervasive, and the (expected) MWV must be defined for the range of storage levels through which storage trajectories may eventually pass. Thus Scott characterises the release/storage choice in terms of a trade-off between a demand curve for release during the period, and a demand curve for storage at the end of the period. The latter may equivalently be thought of as the supply curve for water to be released in the current period, or as defining the MWV as a function of storage. Scott and Read show that, provided demand curves for release can be defined for each period, a variant of dynamic programming can be used to efficiently determine the MWV curve for any period by backward recursion using those demand curves for release, and an assumed end-of-horizon MWV curve.

Experience shows that, while the MWV curve changes substantially over the year, its general shape is quite consistent. When storage is low, the MWV is high, and vice versa, with the curve being relatively flat for most levels of storage, falling to zero at the upper bound, but rising much more steeply as storage decreases towards its lower bound. This reflects the fact that the system is able to cope with a wide variation of inflows, and hence

---

[3] In the absence of storage bounds, if water had a higher value in period *b* than in current period *a*, it would be profitable to hold water over from period *a* to use in period *b*. This would increase the scarcity (and value) of the water available for use in period *a* and decrease the value of water in period *b*. It would be most profitable to keep shifting water between the two periods until the value of the water in the two periods was equal.

storage levels, at moderate cost, unless storage reaches fairly extreme levels, in which case water becomes very valuable as a means of averting a significant probability of shortage. As before, the optimal strategy is to try to keep the (expected) MWV constant over time. But stochasticity in observed inflows makes holding to this optimal storage policy impossible, so that the MWV fluctuates over time, even when the storage bounds are not reached.

This implies changes to market offers, and hence to prices, and our goal is to capture the implied relationship between storage levels and prices. The hypothesis presented in this chapter is that hydro producers in the NZEM use this MWV methodology in their reservoir management. Therefore, the apparent relationship between storage levels and NZEM spot prices should resemble something of what this methodology would predict. Further, incorporating something of the theory of water value calculation into a model for spot prices will increase its ability to model price behaviour over time. Since the MWV curves the generating companies are actually using are unknown, and it is not desirable to embed an optimisation algorithm in our estimation procedure, some heuristic approximations must be employed.

Rather than utilise an exact optimisation approach, we note that the median historic storage trajectory observed in our sample is probably not a bad proxy for the optimal trajectory for which a hydro generator might be expected to aim, under moderate inflow conditions. And, since that median trajectory is well away from the upper and lower storage bounds, it seems reasonable to assume that the MWV along that trajectory will be not too far away from the constant level which such a generator will try to maintain. If inflows are more or less than anticipated, storage levels will move towards either the upper or lower bound. Then the MWV must decrease or increase respectively, thus acting to stabilise the situation by increasing/decreasing release. The optimal trajectory, from any realised storage level, will then tend to parallel the expected trajectory at a higher or lower constant MWV level, before eventually reverting, as inflows revert to normal levels. So it makes sense to define MWV as a function of the "relative storage level" (RSL), defined in terms of deviation from the expected storage trajectory, and of movement toward the limits of the observed storage trajectory distribution. And, from experience, the rate at which the MWV adjusts, and hence the broad shape of the MWV curve between those limits, is fairly consistent over time.

It should be noted that the measure for the MWV used in this thesis is a function of current storage levels, and therefore is fundamentally backward-looking. In reality, the MWV of a firm reflects the expected value of the water held in storage given an expectation of the distribution of future of spot prices, the likelihood of future inflows, potential actions by other market participants, its level of contracted power generation, et cetera. The actions of a firm (and hence the value of the water held in storage) are likely to be quite different if inflows into its reservoir(s) are forecast to be above or below average levels, if a firm is heavily contracted, or if there is some threat of regulation in the immediate future. These factors add to the complexity of the optimisation models required to solve for the MWV, and, ideally, they should be included in as explanatory variables an econometric model of NZEM reservoir levels and spot prices. However, incorporating measures such as rivals' behaviour and future spot prices are beyond the scope of the research in its this thesis, and we illustrate that the empirical evidence supports the hypothesis that the current storage level is the primary driver of the level of spot prices.

4.4 The NZEM storage and price data

As mentioned in the previous chapter, the Haywards node, which most studies of NZEM prices use as their reference node, is at the northern end of the HVDC link from Benmore to Wellington. The initial aim of our study was to use prices from this node as well, and to use aggregate storage data for the whole country. However, the presence of the HVDC link creates a variation in the nodal prices around New Zealand's electricity network. In general, the flow of electricity across the link is from south to north, as the South Island contains the major seasonal storage in New Zealand. When the capacity of the link comes close to being constrained, the prices at either end of the link can become very different from each other. In approximately 3.7% of the half-hourly periods from the start of 2000 to June 2003, the ratio of the price at one end of the link to the price at the other end was greater than 3:2. On one occasion, the price at Benmore was $1/MWh while at the same time in Wellington the price at Haywards was $141/MWh[4].

In order to minimise the effect of the link on prices in the initial stages of this study, final prices from the Benmore market node and the storage level from the Waitaki system (of which the dam at Lake Benmore is a part) are used as indicators of the price level and the storage level. In subsequent work, after illustrating the concept in this chapter, national aggregate storage and prices from the Haywards node are used instead. Explicit modelling of the effect of the HVDC link on price behaviour is left for further research. As can be seen in Figure 4.1 below, which shows the average storage levels over the course of the year, the amount of water held in storage in the Waitaki System is more than is held in the entire rest of the country (storage is divided into four partitions: Waitaki System, the rest of the South Island, Lake Taupo in the North Island, and the rest of the North Island).

[4] See Videbeck (2004) for a thorough discussion of the effects of transmission constraints on nodal prices around New Zealand, and further analysis of the locational differences between NZEM spot prices.

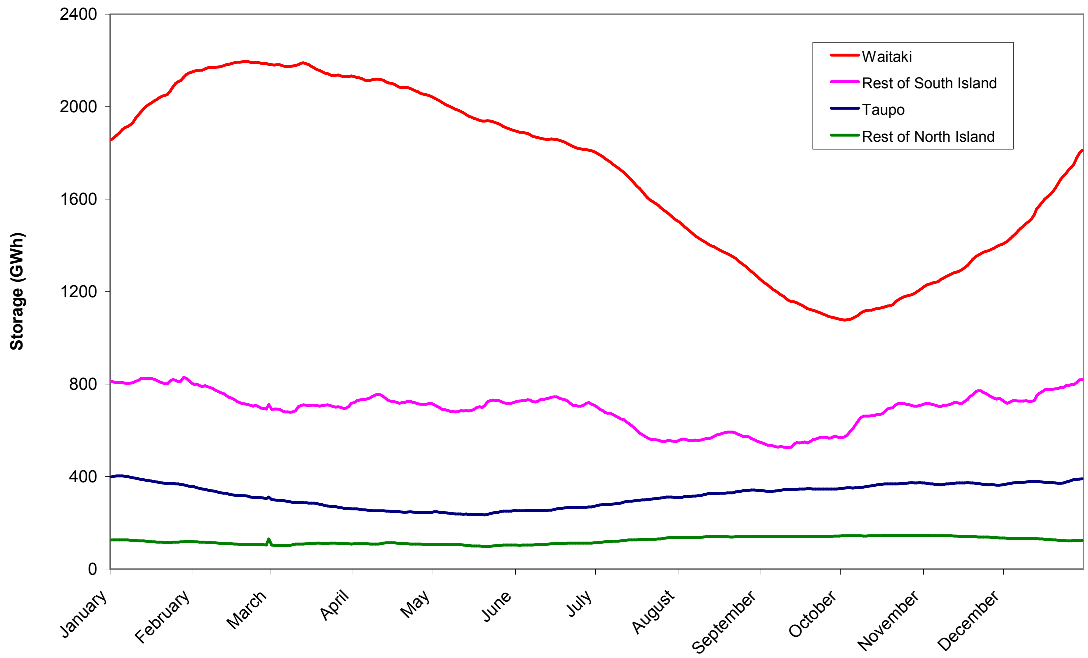

Figure 4.1: Average storage levels in New Zealand for each day of the year, calculated over the period 1980-2003

It is well-known that there is a relationship between the spot price and the storage level. However, as Figure 4.2 shows, there is no apparent relationship between the price and the absolute storage level. Prices have been high when the reservoirs have been both full and empty. Instead, the major hypothesis in this chapter is that the form of the relationship ties in with the MWV theory discussed in the previous section.

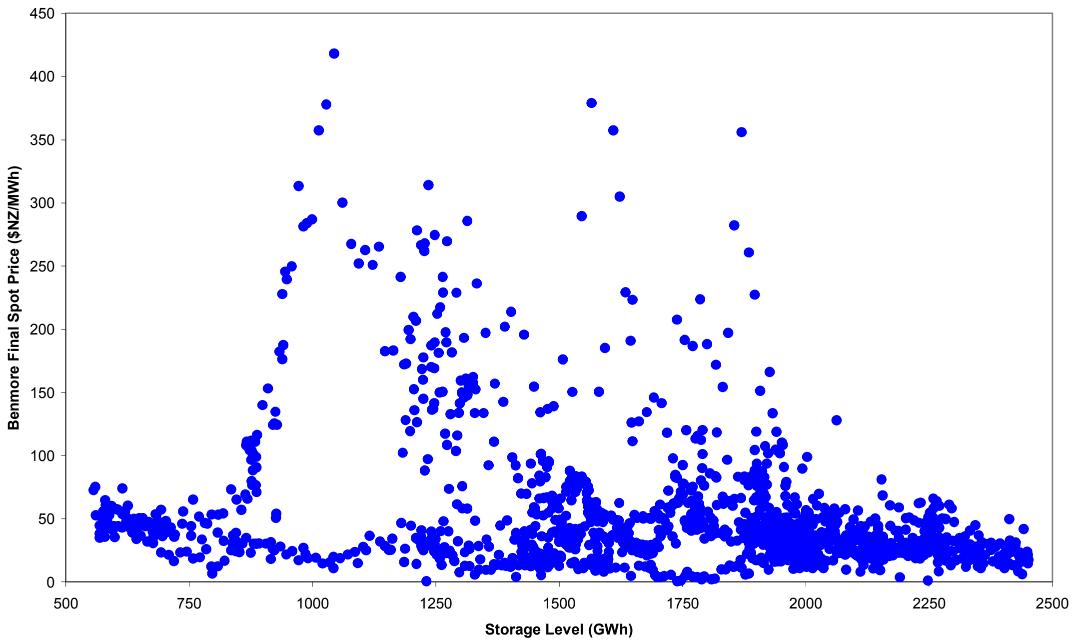

Figure 4.2: Daily average spot prices from the Benmore node versus Waitaki storage level, August 1999 - June 2003

As the storage level approaches its lower bound, the MWV derived by reservoir management models increases at an increasing rate (see Figure 4.5). However, the calculation of the MWV for a certain level of storage depends not just on how much water is held in total, but on how much water the producer would expect to have at that time relative to their planned storage trajectory. For example, the managers of the Waitaki system reservoirs could expect the average storage level trajectory given in Figure 4.1, which would take into account annual average patterns of load, generation and inflows. If their current storage level were higher than the expected trajectory, the MWV would be lower to encourage generation and decrease the storage level. The converse is also true.

To gauge the relative level of storage, it is therefore necessary either to compare the actual storage level with the average storage level, as described above, or to have a lower bound which represents a "danger zone". The closer the storage level gets to the danger zone, the higher the MWV. In this study, the use of two different lower bounds is explored:

- For each day of the year, the lowest storage level that had been recorded between 1980 and 2003
- The historic tenth percentile of storage levels for each day of the year, over the period from 1980 to 2003

Using the lowest storage level proves not to work well, as this level is influenced too greatly by one year in particular, 1992, a year in which an extreme sequence of inflows led to very low storage levels. Therefore, experience suggests using the tenth percentile gives the best indicator for a "danger zone" level, and also seems the most intuitive. In reality, any storage level guide for reservoir management would be a smooth curve, so a 45-day moving average of the tenth percentile is taken as the lower envelope. At the time this study was commenced, the available data included final half-hourly spot prices from the Benmore node from 1 August 1999 to 30 June 2003. The lower envelope (the red line) and the storage level of the Waitaki system (the blue line) over this period are shown in Figure 4.3 below:

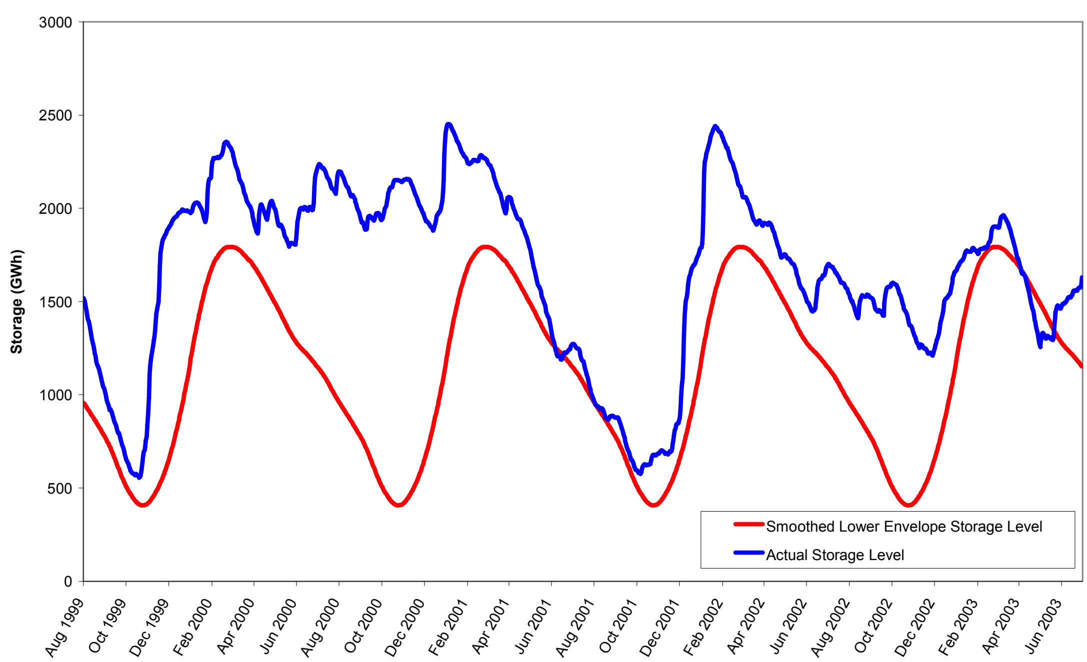

Figure 4.3: Actual storage level and smoothed historic tenth percentile storage level for the Waitaki system, August 1999 - June 2003

The graph above gives a representation of how the storage level has changed between 1999 and 2003. The times of water shortages (which prompted power saving campaigns in New Zealand) are clearly evident in 2001 and 2003. 1999 was also a dry year, however in relative terms it was never as dry as the other two years. Both 2000 and 2002 were wet years, characterised by stable and low prices, which is evident from the graph above.

Given this lower envelope, the next step in the analysis involves establishing what relationship the envelope, the actual storage level and the price actually have. A nett storage position (actual storage less the lower envelope) is calculated to give us the relative storage level (RSL) for each day in the sample period. The RSL on day t is calculated by subtracting the actual storage level from the historic tenth percentile:

$$RSL_t = Storage\ level_t - Historic\ tenth\ percentile_d$$

where d is the day of the year corresponding day t (t=1 on 1 August 1999).

Plotting both time series together as in Figure 4.4 below shows how they are related. The blue line is the time series of daily average spot prices, the heavy green line is the RSL, and the dashed green line represents the RSL corresponding to the smoothed tenth percentile storage level (i.e. RSL=0). When the RSL is high, prices are low, and when it decreases the price tends to increase, as was the case in the first quarter of each of 2001, 2002 and 2003. The extreme high prices of 2001 and 2003 clearly occurred at times when the absolute storage level decreased towards, and even crossed over, the tenth percentile level.

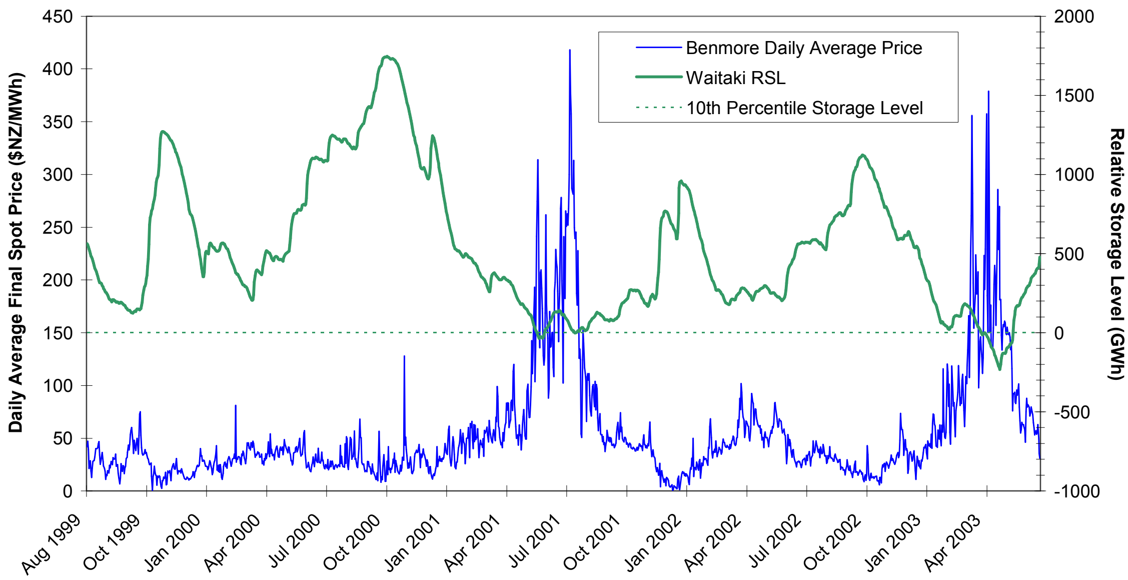

Figure 4.4: Daily average spot prices from the Benmore node, and the relative storage level from the Waitaki System, August 1999 – June 2003

Figure 4.6 plots the RSL against the final spot prices from the Benmore node. Compared to the graph of prices versus actual storage levels in Figure 4.2, the exponential-type relationship is obvious, and the shape of the relationship is clearly similar to that of the theoretical MWV curve, a three-dimensional[5] example of which is shown in Figure 4.5 below.

---

[5] The two dimensions of a MWV curve are storage and MWV. The curve changes over time, adding a third dimension to create a water value surface (WVS), as shown in Figure 4.5.

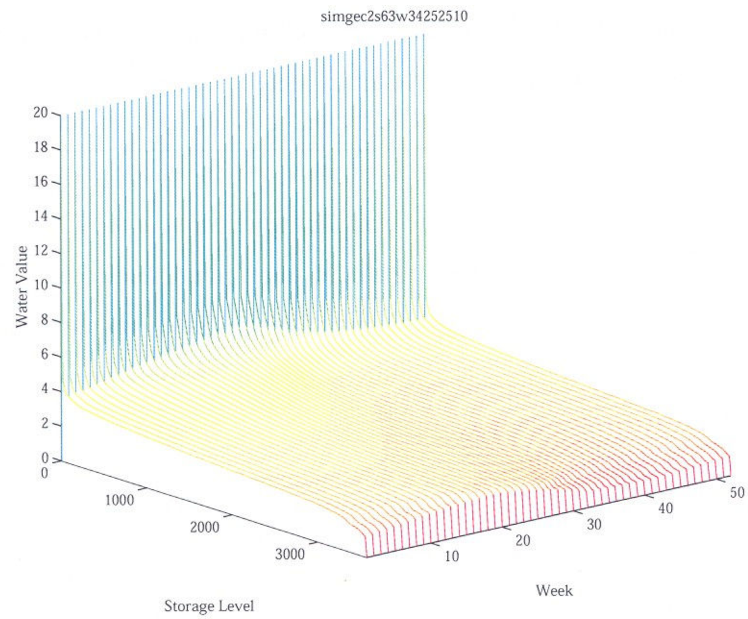

Figure 4.5: A typical MWV surface constructed by a reservoir management model, as shown in Figure 5.5 of Scott (1998) and Figure 8.4 of Batstone (2003)

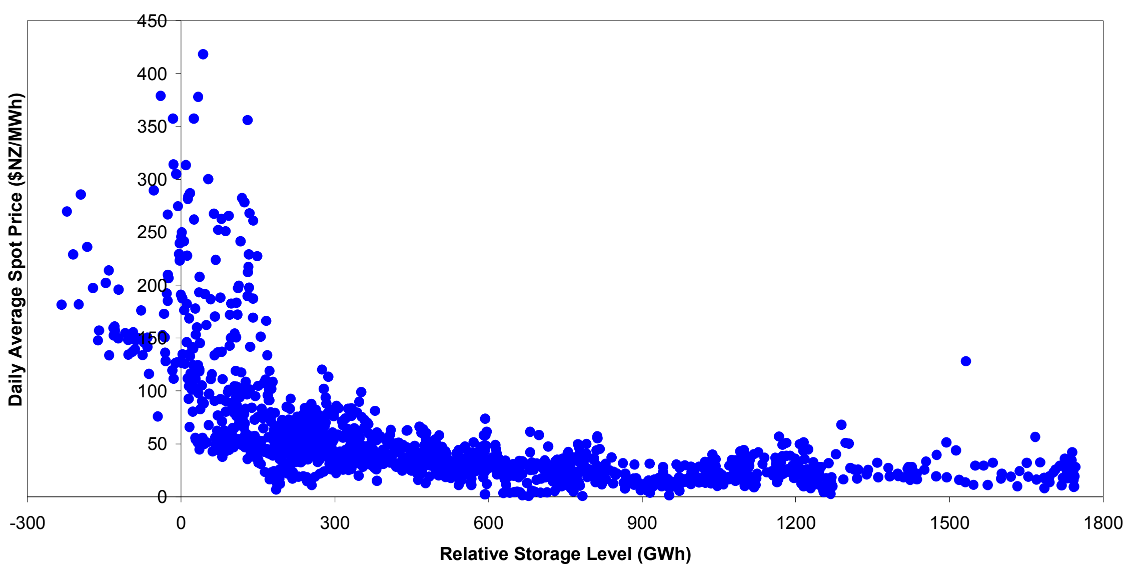

Figure 4.6: Daily average spot prices from the Benmore node versus Waitaki RSL, August 1999 - June 2003

We further hypothesise that the data can be split up into two seasons: a cold season encompassing the sixth months of autumn and winter (March-August) and a warm season including spring and summer (September-February)[6]. Different reservoir management policies may be in place depending on the time of year, and thus the relationship between the RSL and the spot price may not be constant. Demand for electricity peaks during winter in New Zealand, at the same time as inflows decrease due to colder weather. In times leading up to and during the winter we would expect that producers are more conservative in their scheduling to ensure that they have enough water for the peak season.

A graph of average inflows over the course of the year is shown in Figure 4.7. This reinforces the seasonal nature of New Zealand's inflows, with the inflows increasing from (approximately) September, and decreasing from (approximately) February/March. The warm and cold seasons therefore correspond approximately to periods when inflows are expected to increase and decrease respectively.

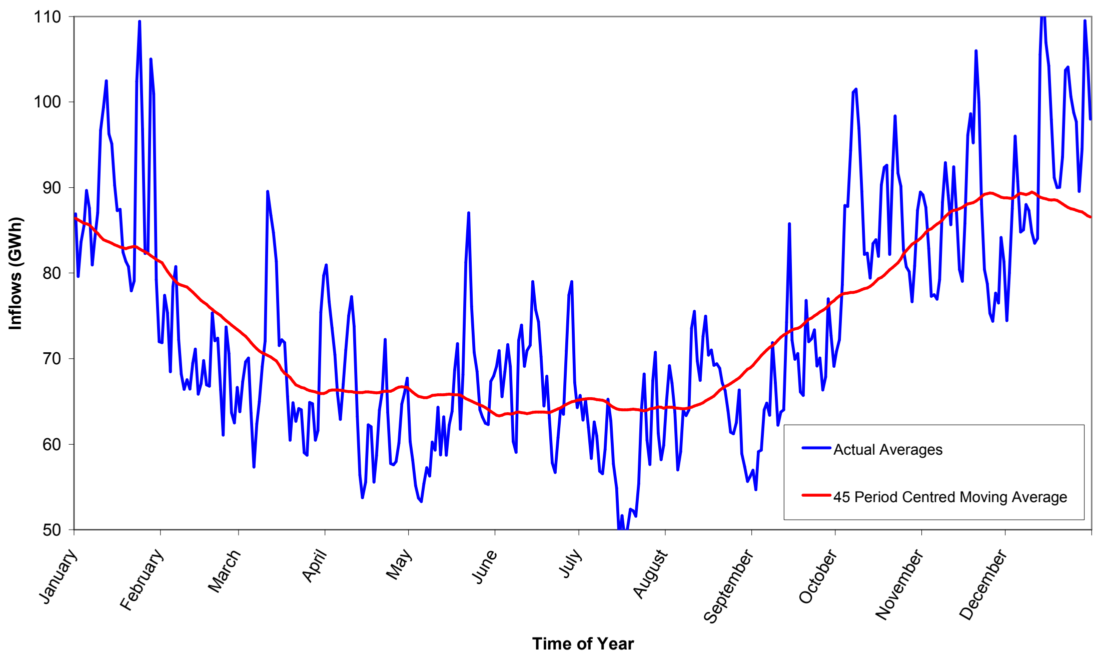

Figure 4.7: Average aggregate inflows into New Zealand hydro lakes, 1980-2003

[6] The timing of these seasons is consistent both with other studies of electricity spot prices and common practice.

Further, spot price behaviour over the sample period is markedly different in the cold season compared with the warm season. Figure 4.8 below shows the monthly average spot prices and monthly standard deviations of daily average spot prices (grouped by month) from August 1999 to June 2003. Prices are clearly higher and more volatile from March to August than from September to February.

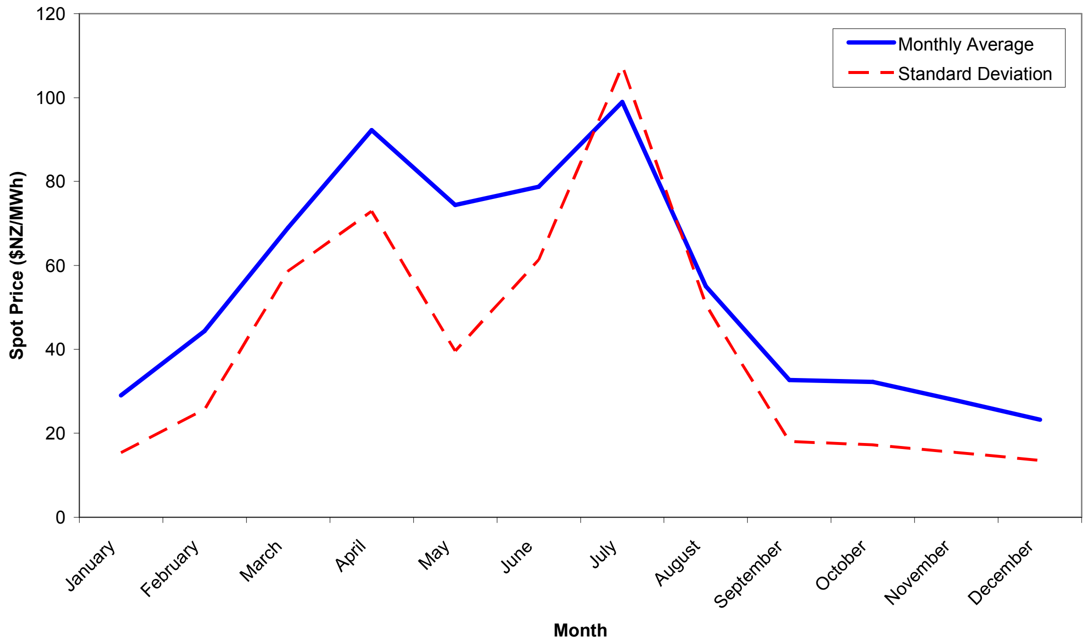

Figure 4.8: Monthly average spot prices (and standard deviations) from the Benmore node, August 1999 - June 2003

Figure 4.7 and Figure 4.8 above suggest that price behaviour (and possibly reservoir management) during the sample period was different in the cold season compared with the warm season. In light of these graphs, and the resulting graphs of the form of Figure 4.6, the decision to partition the data into two seasons seems logical. Splitting the data shown in Figure 4.6 into these two seasons reveals two very interesting graphs. The following charts show the relationships between the RSL and spot prices in the cold and warm seasons respectively:

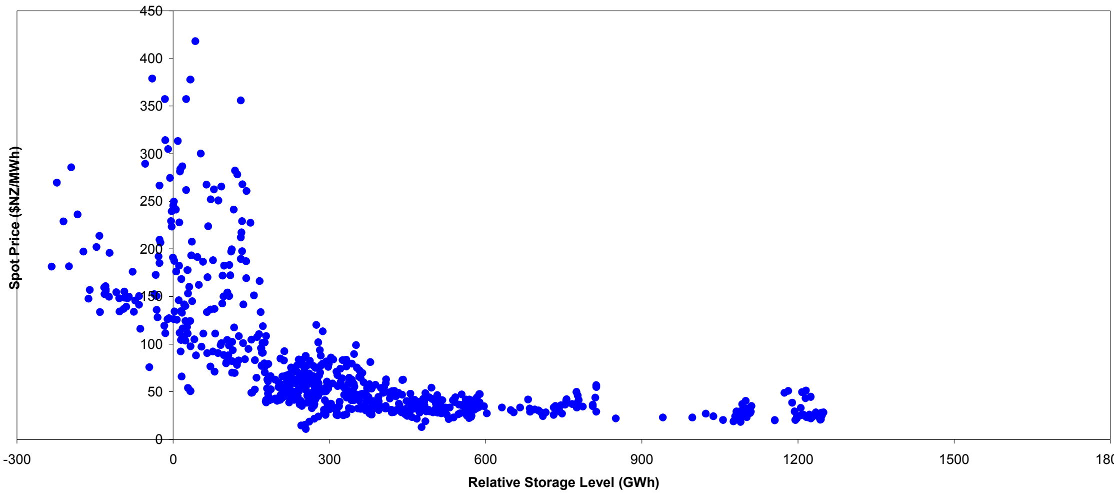

Figure 4.9: Cold Season: Daily average spot prices from the Benmore node versus Waitaki relative storage level, August 1999 - June 2003 (March – August only)

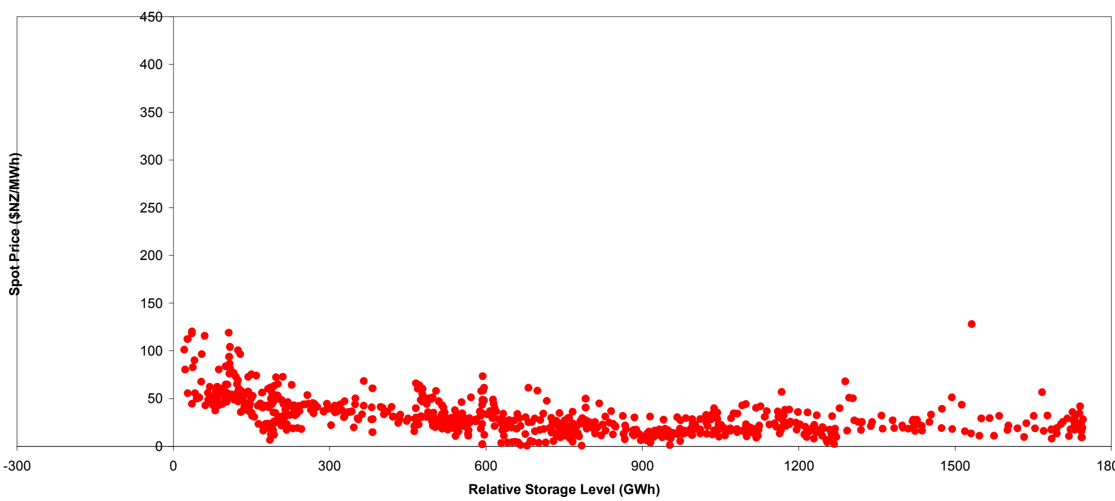

Figure 4.10: Warm Season: Daily average spot prices from the Benmore node versus Waitaki relative storage level, August 1999 - June 2003 (September – February only)

The apparent strength of the relationship between the price and the RSL justifies the inclusion of the MWV function in a model for the spot price. Fitted RSL-price curves can be calculated using an exponential function of the RSL, as an exponential function ties in to the shape of the water value curves discussed earlier and appears to fit the data in Figure 4.9 and Figure 4.10 well. The form of the function used is:

$$MWV_t = c + w\mathrm{e}^{x(y+0.01RSL_{t-1})^z}$$

In this case, only the parameter c, the constant price for very large relative storage levels, has any useful interpretation. The parameter y is required to ensure that the function is still defined for negative relative storage levels, x is required to scale the value of the bracketed terms down to a number that can be raised sensibly to an exponential power, and w is the maximum water value obtainable when the RSL is the greatest negative number the function allows. The RSL is scaled by 0.01 to increase computational accuracy in the procedure used to estimate the parameters of the price model, which is described in the following section.

4.5 Incorporating the MWV function into a top-down model

As illustrated in Chapter 3, the deterministic component of the Escribano, Peña and Villaplana (EPV; 2002) model fits the underlying price movements in the NZEM between 1999 and 2003 very poorly, due to the influence the extreme inflow sequences had on those prices. In contrast, Figure 4.14 shows how well the MWV function fits the same price movements (albeit from a different node in the NZEM). As a result, the deterministic component of the EPV model is replaced with the water value function. Equations [1] and [4] to [5] in Chapter 3 remain the same, whereas equation [2] changes from

$$f(t) = Const + Trend.t + \sum_{i=2}^{12} M_i D_{i,t}^M + wkd.D_t^W \qquad [2]$$

to

$$f(t) = c_j + w_j\,\mathrm{e}^{x_j(y_j+0.01RSL_{t-1})^{z_j}} + wkd.D_t^W \qquad [2a]$$

Instead of $j \in$ {autumn, winter, spring, summer} as in the EPV model, now $j \in$ {cold, warm}.

The full price model for the daily average final prices at the Benmore node is:

$$P_t = f(t) + X_t \qquad [1]$$

$$f(t) = c_j + w_j \, \mathrm{e}^{x_j (y_j + 0.01 RSL_{t-1})^{z_j}} + wkd.D_t^W \qquad [2a]$$

$$X_t = \begin{cases} \theta X_{t-1} + e_t & \textit{with probability } (1-\lambda_j) \quad \textit{No Jump} \\ \theta X_{t-1} + e_t + \mu_j + \sigma_j \varepsilon_{2t} & \textit{with probability } \lambda_j \quad \textit{Jump} \end{cases}$$

where $e_t = \sqrt{h_t}\,\varepsilon_{1t}$, and $\varepsilon_{1t}, \varepsilon_{2t} \sim \mathrm{N}(0,1)$. [5]

and $h_t = \omega + \alpha\, e_{t-1}^2 + \beta h_{t-1}$ [4]

As shown in [1], prices are the sum of two components: a deterministic component, which is an exponential function of the RSL (shown in [2a]) plus a dummy variable for weekdays (as opposed to weekends), and a stochastic component. The stochastic component, shown in [5], is the sum of three terms: a fraction (θ) of the previous day's stochastic component, a GARCH (1,1) error (see [4]), and a Poisson jump process. The parameters of the Poisson jump process are allowed to vary depending on the season j. With probability λ_j there will be a jump on any given day. If there is a jump, its size will be drawn from a normal distribution with mean jump size μ_j and variance σ_j^2. Each of these parameters is explained in more detail in the summary of the EPV model in Chapter 3.

As with the models estimated in Chapter 3, we fit the price model above to spot prices from the Benmore node using the CML procedure in GAUSS 6.0[7]. The modification of

[7] As with the models estimated in Chapter 3, the maximum likelihood procedure is constrained by the fact that all variances must be positive, and all probabilities must be between zero and one. No constraints are

the likelihood function (see Appendix B) is simple, involving a straight swap of the *f(t)* function in [2] with the *f(t)* function in [2a]. To enable comparison between the fitting performances of the two deterministic functions, we also fit the EPV model[8] to the same NZEM spot price time series. Table 4. contains the results of fitting both the EPV model and the modified MWV model to daily average spot prices from the Benmore node from 1 August 1999 to 30 June 2003.

EPV Model		
1 Aug 1999 - 30 June 2003		
	Coefficient	t-stat
CONSTANT	15.94	4.51
TREND	0.0098	2.36
M_2	3.82	1.83
M_3	9.67	3.42
M_4	15.68	4.82
M_5	10.61	2.52
M_6	6.10	1.53
M_7	6.69	1.78
M_8	0.51	0.14
M_9	0.59	0.17
M_{10}	1.35	0.43
M_{11}	0.48	0.16
M_{12}	-1.18	-0.72
WEEKDAY	2.95	10.16
θ	0.92	91.45

EPV Model incl. Water Value		
1 Aug 1999 - 30 June 2003		
	Coefficient	t-stat
c_{COLD}	22.19	991.05
w_{COLD}	481.24	755688.08
x_{COLD}	-0.1519	-2.46
y_{COLD}	7.19	55.02
z_{COLD}	1.34	62.33
c_{WARM}	8.82	347.62
w_{WARM}	115.70	28464.76
x_{WARM}	-0.3845	-19.41
y_{WARM}	2.76	28.39
z_{WARM}	0.78	52.00
WEEKDAY	3.06	32.43
θ	0.86	40.12

placed on any of the values of the MWV function, and no assumptions are made regarding the distributions of these parameters.

[8] As described in Chapter 3.

ω	2.76	3.69
α	0.34	7.34
β	0.58	14.67
λ_{COLD}	7.84%	3.82
μ_{COLD}	17.26	3.18
σ^2_{COLD}	307.79	1.71
λ_{WARM}	6.30%	3.13
μ_{WARM}	10.37	2.26
σ^2_{WARM}	351.33	2.98
Log-Likelihood	-5114	
SIC	10403	

ω	2.84	41.53
α	0.34	24.84
β	0.57	30.79
λ_{COLD}	8.17%	5.19
μ_{COLD}	17.96	2636.38
σ^2_{COLD}	249.45	1372291.34
λ_{WARM}	5.56%	2.60
μ_{WARM}	12.10	1765.66
σ^2_{WARM}	383.02	1707156.66
Log-Likelihood	-5095	
SIC	10342	

Table 4.1: Estimated parameters from application of EPV model and EPV model including two seasonal MWV functions to daily average spot prices from the Benmore node: August 1999 – June 2003

Initial comparisons between the estimated parameters of the EPV model for the time series of prices from the Haywards node (see the results in Chapter 3) and the Benmore node reveal the same seasonal patterns, trend and weekday effect, as you would expect. Interestingly, the estimate of the unconditional variance parameter ω for the Benmore node is just over half the size of that estimated for the Haywards node, suggesting that the time series from Benmore is substantially less volatile[9]. A detailed comparison of estimated parameters between nodes is outside the scope of this study and has been left

[9] This finding is in line with Mason (2002) who reports that, in general, variation in prices is greater at the Haywards node than at Benmore, presumably reflecting the fact that, when the HVDC link is fully loaded, the North Island becomes a thermal-dominated system.

for further research, however Videbeck (2004) presents analysis of the behaviour of prices at different nodes in the NZEM in more detail.

Overall, incorporating the storage level has increased the fit of the model to the historic price data, as can be seen by both the increase in the log-likelihood and a decrease in the Schwartz Information Criterion[10]. However, identifying changes in the individual parameter estimates themselves offers perhaps the best indication of the increase in the fit.

A striking feature of the estimated stochastic parameters is how similar they are in the two models. Most notable in the set of estimated jump parameters is the fact that the standard deviation of the jumps in the cold season is 17.5 in the EPV model, but only 15.8 in the model including the MWV. Combined with the similarity of the other stochastic parameters, this indicates that more of the variation in the price has been accounted for by the MWV deterministic component than in the EPV model.

The most significant difference overall, however, is the decrease in the auto-regression parameter θ. One conclusion that can be drawn from this change is that the underlying mean level of the prices has been modelled more accurately, and therefore reversion to the modelled mean (which, in the revised model, is conditional on the RSL) occurs at a faster rate than in the EPV model. As prices in the NZEM can be driven by the storage level to both low and high levels for sustained periods of time, it is logical that including the driver of these price movements in the model will improve the fit and reduce this parameter estimate.

[10] As described in the previous chapter, given a certain model specification, the CML routine maximises the likelihood of a time series of observations having been produced by a particular set of parameters. The greater the log-likelihood, the higher the probability that the estimated set of parameters is correct, and the better the fit to the data. The Schwartz Information Criterion (SIC) is a function of the log-likelihood, adjusted by the number of parameters in the model. Therefore, if two models yield the same log-likelihood, the model which has fewer parameters will have the lower SIC.

Figure 4.11 illustrates the differences between the two deterministic components fitted to the price series. The straight line with monthly steps (the purple line) is the estimated EPV component, and the red line shadowing the actual prices (the blue line) is the estimated MWV component.

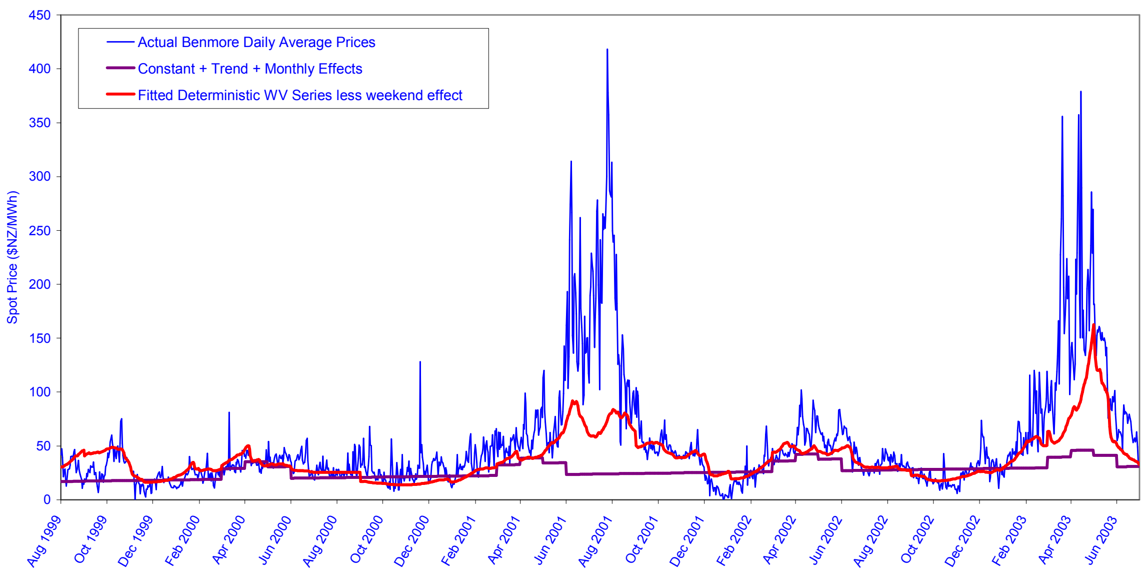

Figure 4.11: Daily average final spot prices from the Benmore node, August 1999 – June 2003, estimated deterministic component (less weekend effect) from the fitted EPV model, and estimated deterministic component (less weekend effect) from the fitted model including two seasonal MWV functions

While the fit of the underlying price level is obviously greatly improved in the two dry years, the fit of the price level in the normal years (2002 in particular) is also very accurate. Interestingly, the estimated MWV functions do not appear to fit the dry year prices as well as in Figure 4.14. In Figure 4.14, the prices in May 2003 were mostly overestimated, creating a number of large negative residuals. In contrast, the MWV function in Figure 4.11 largely underestimates the prices in the dry years of 2001 and 2003, as the large positive residuals are accounted for by the jump distribution.

Even though the formal estimation procedure provides an improved fit to the price series, it is worth noting that the visual fit of the MWV function to the spot price time series can be improved further by fitting the parameters of the MWV function by eye. The following graphs illustrate the fit obtained by fitting the two seasonal MWV functions by eye, and are included in this thesis purely as an illustration of concept.

The following water value functions are fitted by hand to the cold and warm seasons respectively[11]:

$$\textit{Cold Season MWV} = 23.5 + 390\,e^{-0.2725(2.5+0.01RSL)^{1.33}} \qquad r^2 = 45.76\%$$

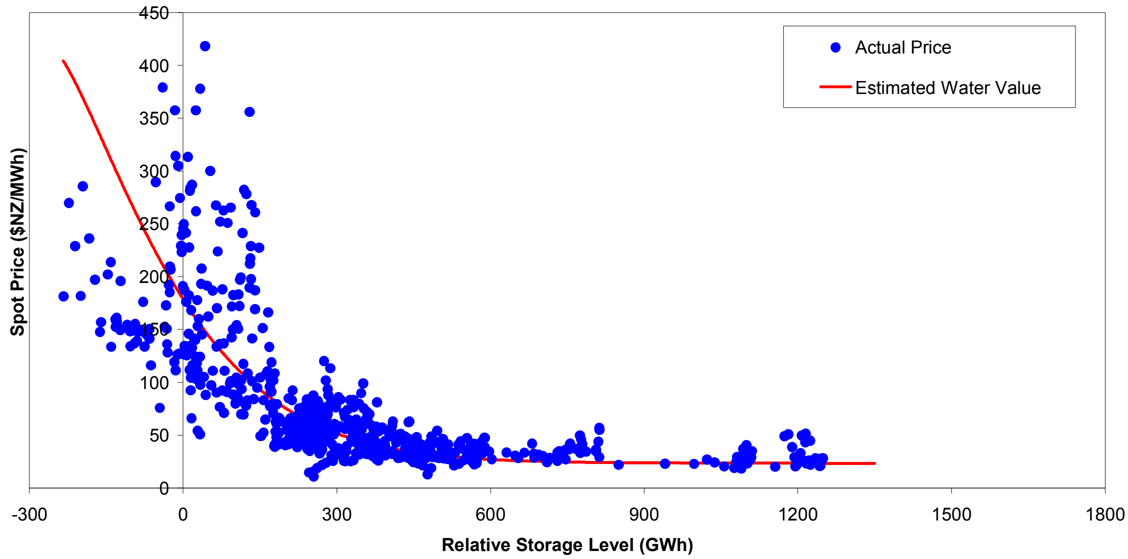

Figure 4.12: Cold Season: Daily average spot prices from the Benmore node versus Waitaki relative storage level, and fitted MWV function fitted by eye, August 1999 - June 2003 (March – August only)

[11] Despite the fact that these fits appear accurate, when the parameter estimation search procedure in GAUSS is given these parameter values as a starting solution, the search moves away from these values to the values given in Table 4.. It is possible that a different solution method may have yielded different results, however as the procedure used in this thesis is the CML method, all fits are assessed on the values of the log-likelihood function and functions thereof. As with the modelling presented subsequently in Chapter 8, the use of different measures for goodness of fit could be explored in future research.

$$Warm\ Season\ MWV = -10 + 275\,e^{-0.58(5+0.01RSL)^{0.50}} \qquad r^2 = 33.47\%$$

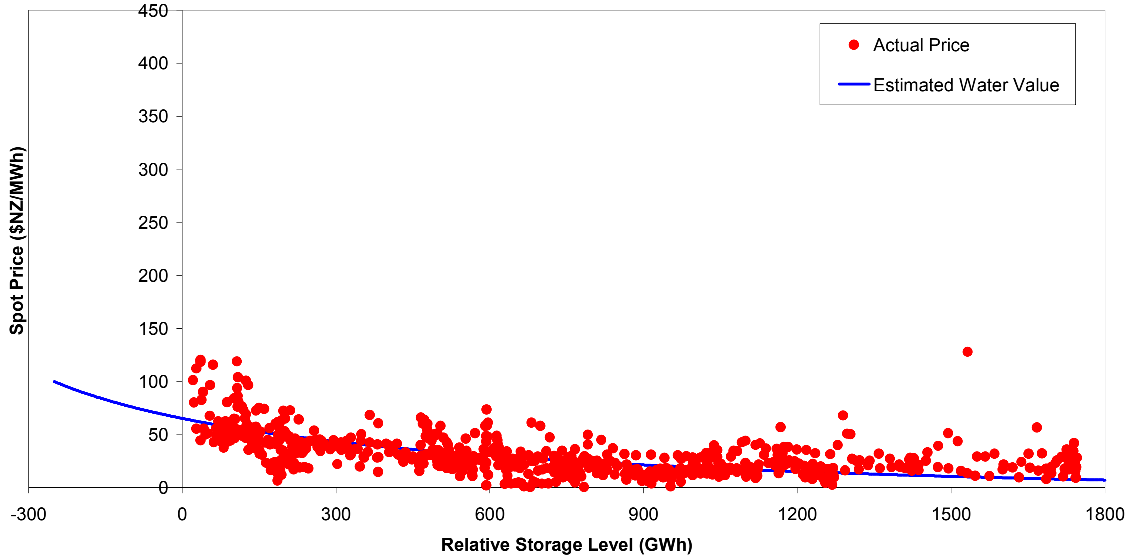

Figure 4.13: Warm Season: Daily average spot prices from the Benmore node versus Waitaki relative storage level, and MWV function fitted by eye, August 1999 - June 2003 (September – February only)

Assuming that prices over the length of the sample period can be modelled using these two deterministic functions alone (with no stochastic process at all) provides the following fit to New Zealand price data, which appears to model actual price behaviour very closely:

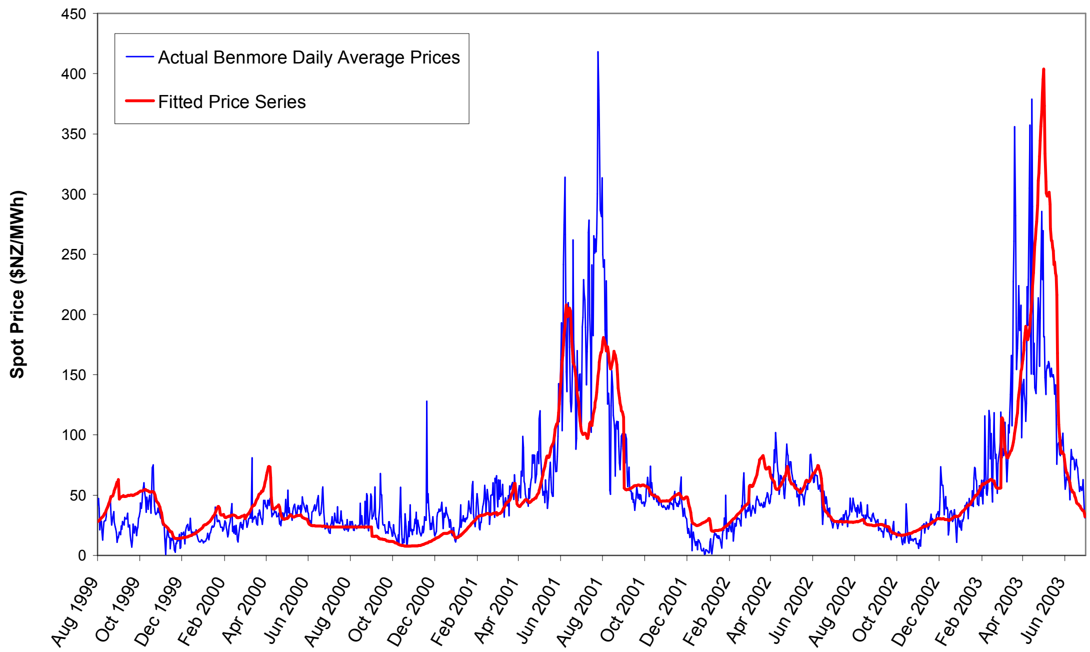

Figure 4.14: Daily average final spot prices from the Benmore node, August 1999 – June 2003, and fitted MWV function based on the relative storage level

Using these two MWV functions in a forecasting framework (again, with no stochastic process) also yields very close fits between the price duration curves[12] of the actual and fitted data. The water value functions alone provide a very tight fit to the New Zealand daily average prices. While it does not show how close the fit is on a day-by-day basis, the PDC in Figure 4.15 shows the closeness of the fit on an aggregate basis.

[12] A price duration curve (PDC) is essentially the complement of a cumulative distribution function – it shows the proportion of time that the price is above a certain level, rather than below.

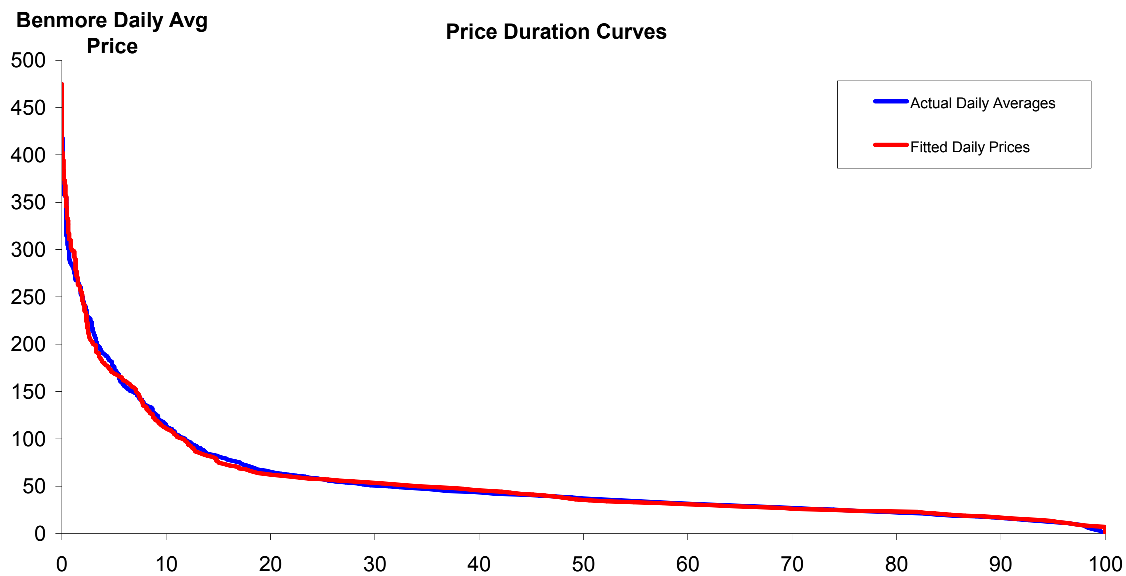

Figure 4.15: Price duration curves for daily average spot prices from the Benmore node, August 1999 – June 2003: Actual prices and fitted MWV

4.6 Conclusions

Two major conclusions can be drawn from the research presented in this chapter.

Firstly, it has been shown that there is a clear relationship between the relative storage level (RSL) and the spot price, which is consistent with hydro reservoir management theory. As storage decreases relative to the "danger zone" for that particular time of year, the spot price increases, giving a signal to the market that the security of supply through hydro generation is decreasing. The evidence is consistent with the hypothesis presented in this thesis that hydro generating companies are using the marginal water value (MWV) concept explained in Section 4.3 to value the water they have in storage[13].

Information regarding the reservoir management strategies of hydro generating firms is confidential; therefore we cannot model these strategies explicitly. However, the clarity

[13] While there does appear to be a relationship between the RSL and spot prices, this does not necessarily rule out that hydro generating companies may be using other methods to value the water they have in storage.

of the RSL-price relationship suggests that a function of the RSL can be used to calculate the underlying level of the spot price. This then enables price behaviour to be modelled in accordance with what *actually happens* in the market, rather than operating a bottom-up optimisation model under certain assumptions to estimate what *might* be happening.

Secondly, low relative storage levels result in periods of sustained high prices in the NZEM, which can only be modelled effectively if the storage level is incorporated somehow. Including and estimating a function for the estimated MWV in a sophisticated spot price model improves the fit of the model to the NZEM spot prices, and the estimated parameters of that MWV function are statistically significant. When plotted alongside the deterministic fit given by a linear step function regression, the increased explanatory performance of the MWV function is clearly evident. Also, less of the volatility in the spot price is required to be modelled by the stochastic parameters in the model when a function for the estimated MWV is included, meaning these parameters give a more accurate representation of the true stochasticity in the spot price.

As explained in the previous chapter, it has long been understood that there is a relationship between the hydro storage level and the price of electricity in New Zealand. Since the advent of electricity market deregulation and the formation of the NZEM, this relationship has led to at least two periods of sustained high prices due to water shortages. The research presented in this chapter provides the rationale behind this relationship, illustrates the effect that the storage level has on spot prices, and establishes a method whereby the relationship can be quantified for the purposes of spot price modelling.

5

EXTENDING THE NZEM PRICE MODEL

The analysis presented in the previous chapters showed that a major driver of the level of spot prices in the NZEM is the hydro storage level. The fit of a top-down model to a time series of NZEM spot prices can be improved substantially by incorporating a feature from bottom-up optimisation models of mixed hydro-thermal systems, the marginal water value (MWV). The MWV is a measure for the marginal cost of hydro generation, which is the dominant form of generation in the NZEM. The research presented is consistent with the hypothesis that New Zealand's hydro generators appear to be using the MWV concept in their reservoir management, and illustrates that a simple function of the relative storage level (RSL), a concept introduced in the previous chapter, can track the underlying level of spot prices remarkably well.

The price model presented can be extended in many ways to test various hypotheses regarding NZEM price behaviour. For example, in the last chapter a method for calculating of the RSL was presented. This method involves subtracting from the current storage level a lower envelope storage level, given by a smooth function of the historic tenth percentile storage level. The lower envelope signifies a "danger zone" for the storage level – when the storage level is outside that zone, the RSL is positive and the MWV is lower. However, when the absolute storage level is inside the danger zone the RSL is negative, which leads to a high MWV. One extension for the price model presented in this chapter is allowing the envelope to gradually increase or decrease in order to assess whether reservoir managers are becoming more or less conservative in the light of increasing load.

It makes sense also to assess whether or not the stochastic component in the price model can be linked to the MWV, in order to explain a greater proportion of the volatility in prices not explained by the deterministic component. Using available generating capacity as a measure of both the probability and the likely size of jumps in the price series has been suggested in the academic literature, however the RSL could be the key to modelling price volatility in markets like the NZEM where hydro is the major form of generation. With a storage-induced reduction in available capacity, the offer stacks of hydro generators are likely either to be severely truncated or become much steeper, which would lead to more volatile prices from changes in demand.

This chapter presents extensions to both the deterministic and stochastic components of the NZEM price model, with the rationale behind each extension explained. The final model is presented at the conclusion of the chapter, however the base model for which these extensions are considered is described in equations [1a], [2b], [4] and [5] below, as in the previous chapter:

$$P_t = MWV_t + X_t \tag{1a}$$

$$MWV_t = c_j + w_j \, \mathrm{e}^{x_j (y_j + 0.01 RSL_{t-1})^{z_j}} + wkd.D_t^W \tag{2b}$$

$$j \in \{\text{cold, warm}\}$$

$$X_t = \begin{cases} \theta X_{t-1} + e_t & \textit{with probability } (1-\lambda_j) \quad \textit{No Jump} \\ \theta X_{t-1} + e_t + \mu_j + \sigma_j \varepsilon_{2t} & \textit{with probability } \lambda_j \quad \textit{Jump} \end{cases}$$

where $e_t = \sqrt{h_t}\,\varepsilon_{1t}$, and $\varepsilon_{1t}, \varepsilon_{2t} \sim \mathrm{N}(0,1)$. [5]

and $h_t = \varpi + \alpha\, e_{t-1}^2 + \beta h_{t-1}$ [4]

5.1 Extending the deterministic component of the price model

Graphs in the previous chapter illustrated the closeness of the fit of the estimated MWV to the underlying price level. As a result of this, only one extension to the deterministic component of the price model is presented in this chapter. Other potential calculations of the MWV are presented in Appendix J.

5.1.1 Representing an increase in demand over time

Demand for electricity in New Zealand increases at approximately 2% per year[1]. Using the RSL methodology, one way to test the hypothesis that reservoir management is changing over time in the face of the increasing demand is to determine whether or not the lower storage envelope or "danger zone" is increasing over time. As generating capacity on the Waitaki River (and hydro generating capacity in general) has not increased by any appreciable amount over the period of this study, the increase in demand

[1] http://www.electricitycommission.govt.nz

may have lead reservoir managers to become more conservative in their behaviour, and thus assign higher values to the same relative levels of storage.

To test this hypothesis, we allow the lower storage envelope (currently the tenth percentile of historic storage levels) to vary over time according to a linear trend, which may be either upward- or downward-sloping. In the basic model, the RSL at day t is calculated as:

$$RSL_t = Storage\ level_t - Historic\ tenth\ percentile_d$$

where d is the day of the year corresponding day t (t=1 on 1 August 1999). However, with the lower envelope being able to increase or decrease over time, the RSL is now calculated as:

$$RSL_t = Storage\ level_{t-1} - (Historic\ tenth\ percentile_d + \delta t)$$

where δ is the extra parameter in the model requiring estimation. This new RSL_t is included in equation [2b] of the model as before, with equations [1a], [4] and [5] remaining the same. The model is estimated using the Waitaki storage level and Benmore spot prices as before.

The estimated parameters for this model are shown in Appendix C. The most interesting result is the positive estimated value for the δ parameter, which supports the idea that reservoir managers have become more conservative over time to cope with the increased demand. The increasing lower envelope is plotted in Figure 5.1 below. The dotted red line represents the original case where $\delta = 0$, and the solid red line where $\delta = 0.22214$ as in Appendix C. Another observation that can be made from Figure 5.1 is that the increasing trend could explain why prices were not appreciably high in 1999 despite the relatively low level of storage.

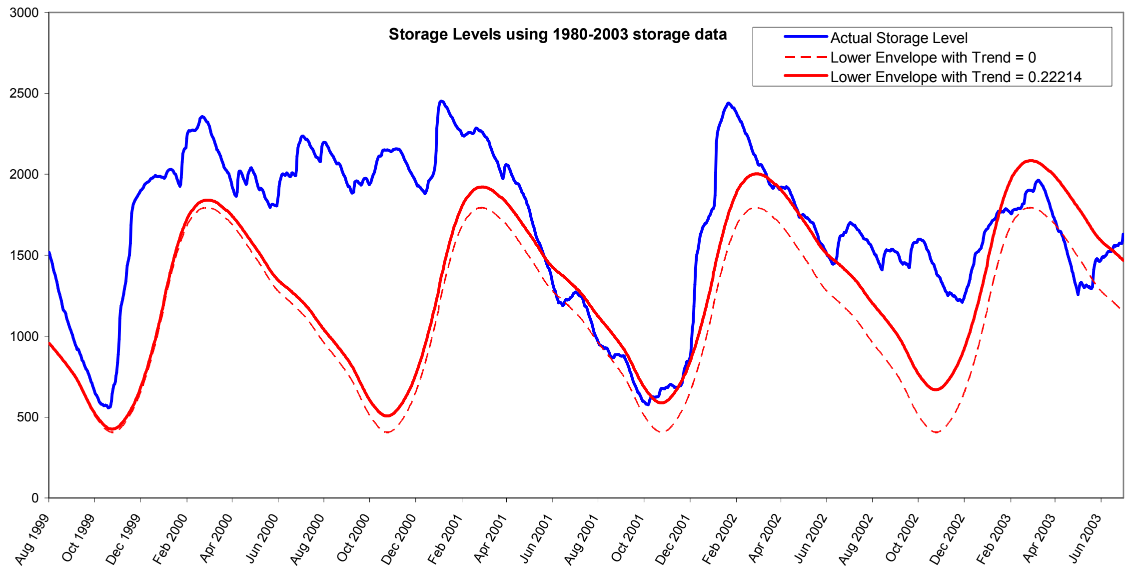

Figure 5.1: Actual storage level, smoothed tenth percentile storage level and trended smoothed tenth percentile level for the Waitaki system, August 1999 - June 2003

The values of the other parameters estimated for this model do not differ markedly from those in the model estimated in the previous chapter. However, despite the fact that the log-likelihood and SIC indicate an improved fit to the price data using the linearly-trended envelope, the estimated value for δ and the estimated values for two of the MWV function parameters are not statistically significant at the 5% level. Therefore, the trended envelope is not included in any further analysis.

5.2 Extending the stochastic component of the price model

As mentioned above, incorporating the MWV into the stochastic component of the price model as a driver of price volatility is a worthwhile extension with an intuitive rationale.

The RSL takes into account both seasonal variations in load and inflows in its calculation. The MWV, calculated as a function of the RSL, reflects the short- to medium-term availability of hydro generation in the market – if the value is high, the level of storage is relatively low for that particular time of year, and vice versa. The greater the MWV, the greater the cost of releasing an extra unit of water now rather than storing it for later release, and the greater the risk of running out of water later in the year given expected inflow and load requirements.

As explained in Chapter 2, electricity prices exhibit a large degree of price-dependent volatility, due to the convex and step-shaped nature of the industry supply curve. In other words, the higher the level of the spot price, the more volatile it will be. When demand for power is low, for example in summer nights in New Zealand, only the “must-run” and cheapest base-load generation needs to be dispatched to meet that demand. Small increases in demand will not likely require much more expensive generation to come online, as the base-load generation units are generally large. However, when demand is high (such as occurs in the winter mornings and evenings in New Zealand), more expensive generation technology (often in smaller units) must be dispatched, and the demand curve intersects the market supply curve at a steeper part of that curve. Small fluctuations in demand will then lead to much larger changes in the price level as the marginal unit is more likely to change, and the vertical distance between steps in the supply curve is greater than for low levels of demand.

That the NZEM spot price can be modelled as a decreasing function of the RSL has already been established. Therefore, the supposition could be made that volatility in the price also depends on the storage level – the lower the RSL, the more volatile the price. The intuition behind this idea is simple – if the RSL is low, the MWV, which is the cost of fuel for the hydro generator, will be high. In the absence of contracting, the hydro generator’s offer stack will shift upwards and become steeper, which has the effect of truncating the market offer stack and making it steeper. Units of hydro that were previously dispatched more often, and for a low price, may now be dispatched only infrequently for a much higher price. Any time the offer stack becomes truncated and steeper, price volatility is likely to increase.

Evidence of the link between the RSL and the volatility of spot prices within each day is shown in Figure 5.2 below, which plots the RSL against standard deviation of half-hourly prices within each day, for each day in the sample period. A similar exponential-type relationship to that shown between the RSL and daily average spot prices is evident. The majority of the days in the sample period that have a high intra-daily variation in prices (i.e. standard deviation of the prices within the day exceeds $50/MWh) also have a low RSL, suggesting the low RSL may indeed be a driver of price volatility within days.

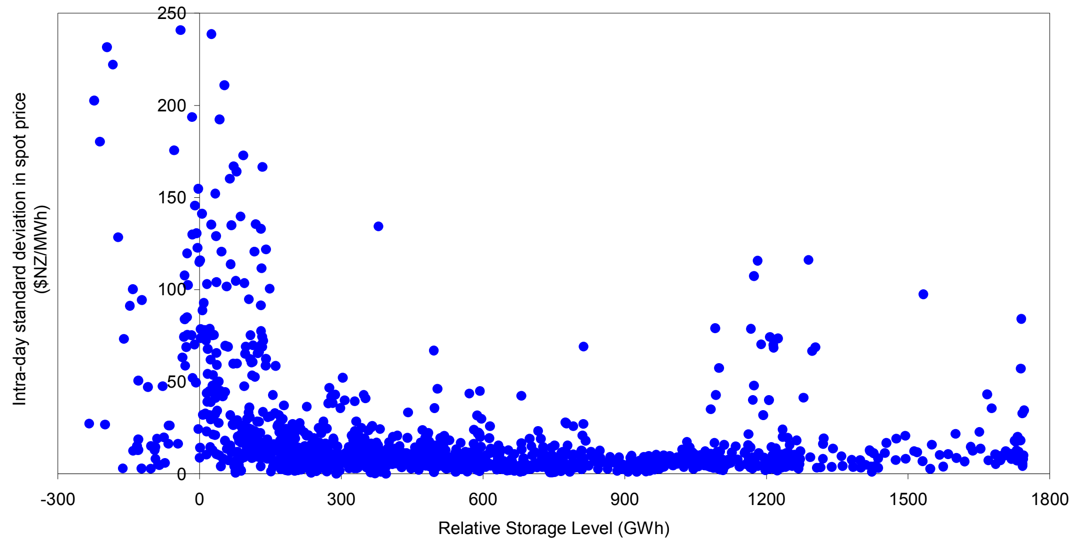

Figure 5.2: Daily intra-day standard deviation of half-hourly spot prices (Benmore node) versus Waitaki RSL, August 1999 - June 2003

This thesis does not examine intra-daily price volatility in any great detail, instead focussing on modelling the level and change in prices from one day to the next. A more relevant graph for the purposes of this thesis is shown below in Figure 5.3. This shows the absolute value of the change between daily average spot prices from one day to the next, plotted against the RSL for the next day. Clearly, there appears to be a relationship between the magnitude of the change in price from one day to the next and the RSL. Apart from a single spike on 19 November 2000, which occurred when the RSL was high, the RSL was low for all other times in which the price changed substantially from one day to the next. It is this relationship that motivates the extensions to the price model explored in the rest of this chapter.

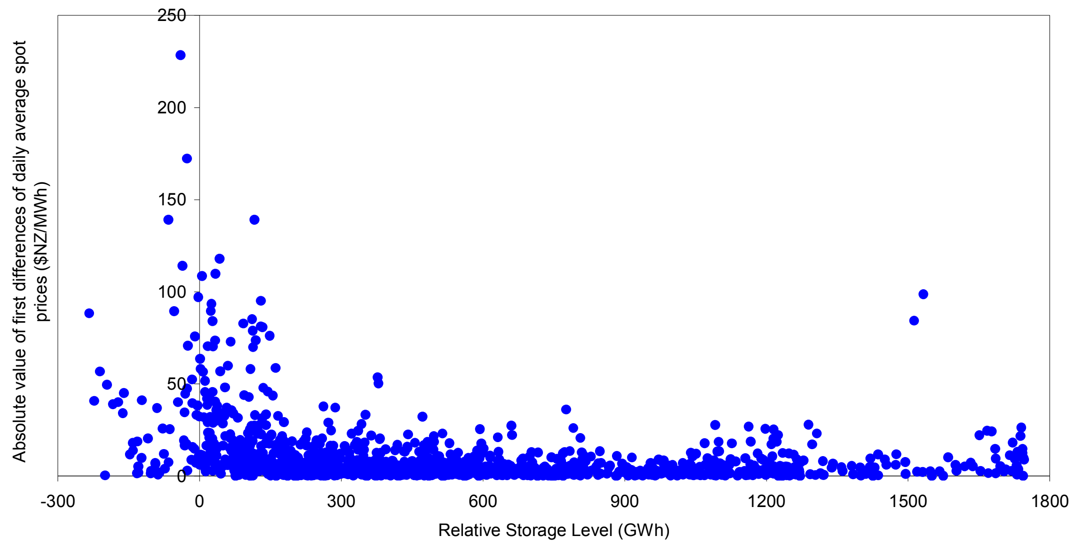

Figure 5.3: Absolute value of first differences between daily average spot prices on adjacent days (Benmore node) versus Waitaki RSL, August 1999 - June 2003

5.2.1 Monte Carlo simulation using the price model

In this thesis, spot prices are decomposed as the sum of two components – deterministic and stochastic. When forecasts of the price are required, forecasts of each of the two components must be calculated first. Forecasting the deterministic component is straightforward – it is calculated simply and exactly as a function of a known input, the RSL. However, assessing the likely behaviour of the stochastic component requires a method of simulation known as Monte Carlo simulation.

Because there are many inputs to the stochastic component, such as the random shocks in the GARCH process, the occurrence of jumps and the size of these jumps when they occur, there are countless combinations of the realisations of each input that could occur on any given day. Accordingly, an infinite number of prices could be simulated on any day. For each day in a Monte Carlo simulation, a realisation of each stochastic variable is drawn from that variable's distribution, and all the variables are combined together to calculate the stochastic component. This component is then added to the deterministic component and a forecasted price results. The price for the following day is then calculated by drawing new realisations of each variable and combining them together.

This process is repeated for each day in the sample period until a full time series of simulated prices has been calculated. Using the same process, the Monte Carlo method requires a large number of time series to be simulated (usually over 1000), and the series are then aggregated, resulting in 1000 simulated prices for the first day, 1000 for the second, and so on. In this way, instead of forecasting an exact price for each day in the sample period, the method forecasts a distribution of prices for each day. A price duration curve (PDC) can then be used to examine the fit of the distribution of all the simulated prices to the distribution of all the observed prices. When comparisons using PDCs are drawn, the time series aspects of the simulation are ignored.

We use Monte Carlo simulation in this chapter to backcast prices from 1 August 1999 to 30 June 2003. We simulate the 1440 prices in the period 5000 times, which provides a median simulated price[2] along with a 95% simulation interval for each day[3]. This enables the time series behaviour of prices produced by the model to be examined more closely.

Using the price model presented in Chapter 3 and the price model specified at the start of this chapter, Monte Carlo simulations of the price time series over the time period August 1999 – June 2003 provide the overall price duration curve (PDC) in Figure 5.4 below.

[2] The *median* simulated price is presented for each day, rather than the *mean* simulated price, to avoid the results presented being influenced too heavily by large positive outliers, which will occur given that the majority of jumps simulated are positive. The median simulated price should therefore track the underlying level of the actual price better than the mean simulated price, which would be biased upwards.

[3] This interval is included to illustrate the range in which 95% of simulated prices fall.

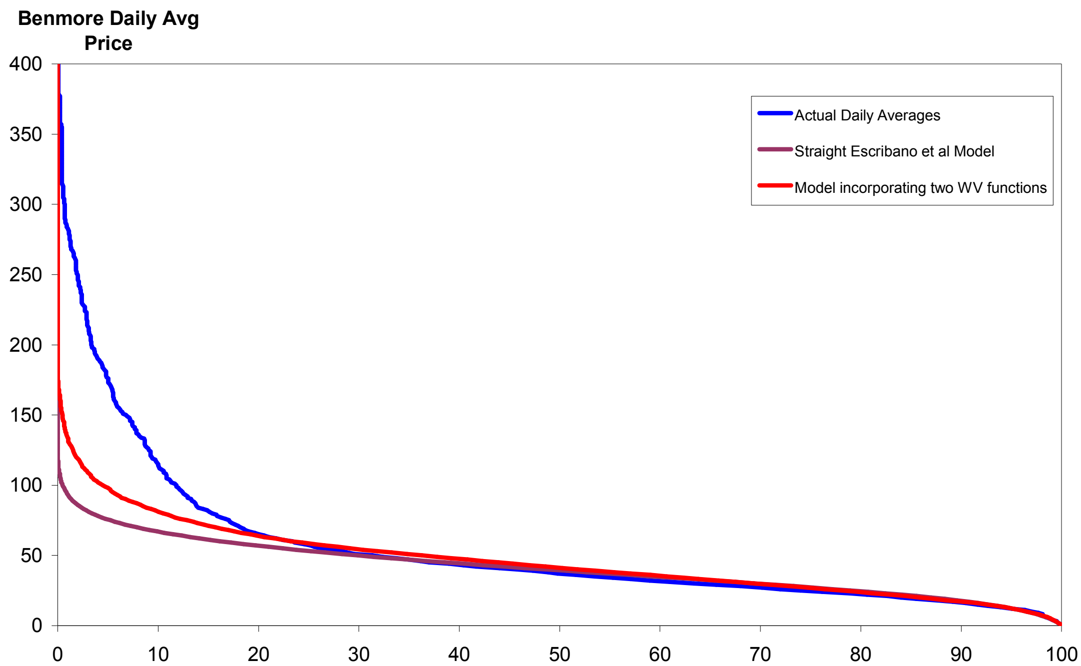

Figure 5.4: Price duration curves for daily average spot prices from the Benmore node, August 1999 – June 2003: Actual prices, prices simulated using the original EPV model of Chapter 3, and prices simulated using the MWV function price model of Chapter 4

This shows that the model including the MWV provides a reasonably accurate fit for the lower 80% of prices, while the EPV model provides a good fit for the lower 70% of prices. However, despite the incorporation of the MWV function, the new model underestimates both the frequency and the magnitude of daily average prices greater than around \$70/MWh. The ability to estimate this top end of the PDC is indeed a vital requirement of any price model, as it is this portion of the graph that is crucial to the valuation of certain financial derivatives and peaking plant. While it is obvious that the model incorporating the MWV function performs markedly better than the EPV model in this respect, it is still far from perfect from a practical point of view.

The PDC shows how well a price model can estimate prices in a time-independent setting, as it is simply an indication of the simulated distribution of prices. However, a perfect PDC of simulated prices could be simulated simply by drawing numbers out of the actual price distribution at random, with no regard for the time processes inherent in the spot price time series, such as jumps and persistence in the volatility.

The following two figures were produced using Monte Carlo simulation to give the median simulated price for each day in the sample period (red line) and the prices at either end of a simulated 95% prediction interval (orange lines). The model in question is the model specified at the start of this chapter and at the end of the previous chapter. The first figure, Figure 5.5, shows the simulation results using a price model excluding the Poisson jump process, whereas Figure 5.6 shows the results using the full model including the jump process.

Both Figure 5.5 and Figure 5.6 show that the median simulated price level for each day of the sample is a reasonable match for most of the observed price time series. However, the two periods when the fit is not good, between June and August 2001 and February and May 2003, occur exactly when the prices increased due to the relatively low storage levels. That the CML procedure calculated estimates that would result in this kind of error is a concern, however exploring exactly why this happened is outside the scope of this research[4].

---

[4] It has become evident through correspondence with other researchers that the CML procedure often produces parameter estimates that, when they are used for forecasting, underestimate the frequency and magnitude of high prices. Bunn and Karakatsani (2003) note that, when using such methods, "Estimation bias is inevitable, as Maximum Likelihood methods tend to capture the smallest and most frequent jumps in the data." As Chan and Gray (2005) note, this tends to result in mean jump sizes and variance being underestimated, but the jump intensity being overestimated.

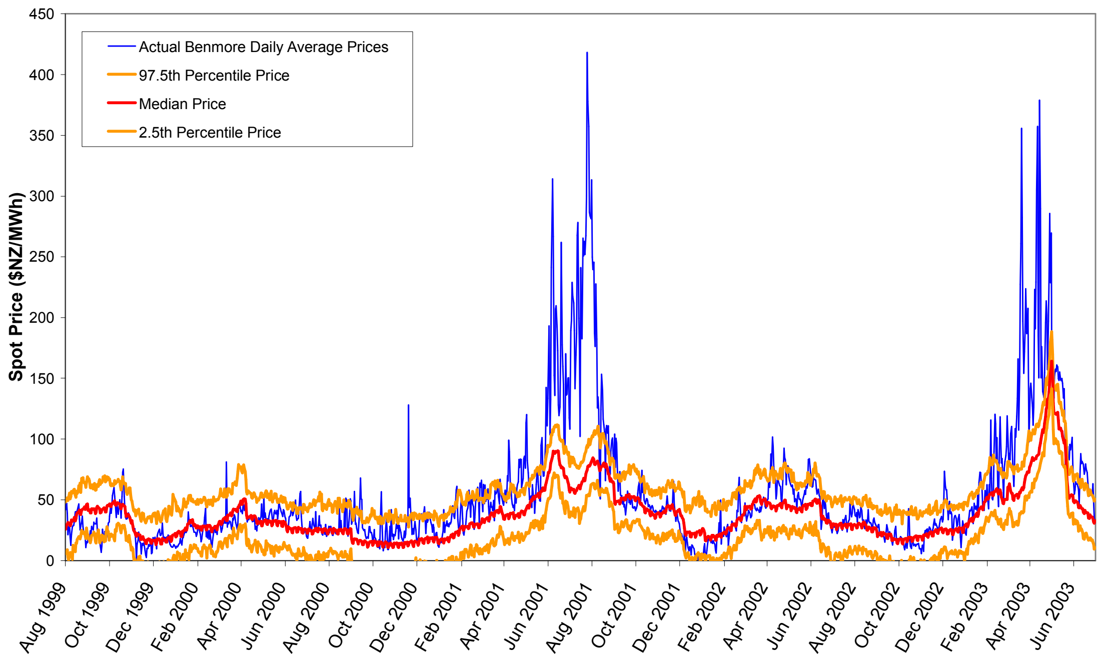

Figure 5.5: **Daily average spot prices from the Benmore node, August 1999 – June 2003 and median, 97.5th percentile and 2.5th percentile simulated prices from running a Monte Carlo simulation over the same period using the MWV function price model but excluding the jump process**

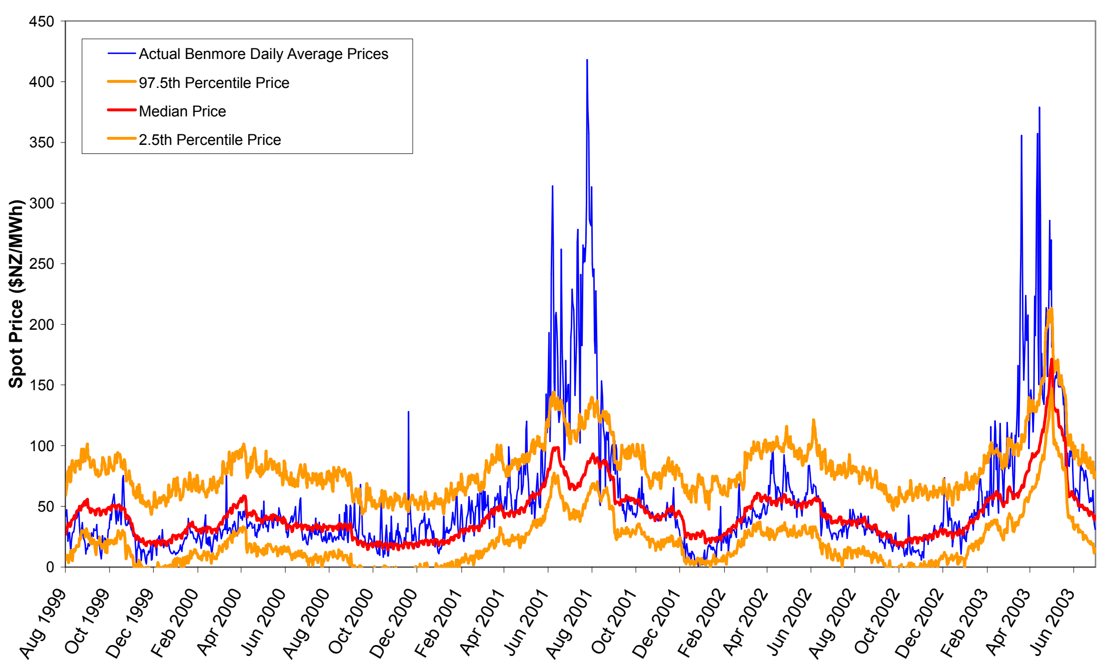

Figure 5.6: **Daily average spot prices from the Benmore node, August 1999 – June 2003 and median, 97.5th percentile and 2.5th percentile simulated prices from running a Monte Carlo simulation over the same period using the complete MWV function price model**

Other observations that can be made from the Monte Carlo simulation results are:

- Without the jump process included, the median price level appears to underestimate the actual price level, on average. The simulation intervals are also much too narrow.
- The prediction intervals and median price level for the full model appear to mimic the underlying movements in the price series remarkably well.
- The simulation interval for the model without the Poisson jump process is symmetric, as expected, in contrast to that of the full model. This reflects the fact that the majority of the simulated jumps are positive.
- The bulk of the actual price series falls within the simulated 95% prediction interval of the full model.

Further to the final observation above, it is worth noting that 11.7% of the actual prices fall outside the 95% prediction interval shown in Figure 5.6. However, if the periods between June and August 2001, and February and May 2003, are excluded from this calculation, this number falls to 3.9%. While excluding these two periods is certainly not valid, and in essence admits that this model cannot simulate prices well in the dry periods, we believe that this fault is not necessarily a conceptual one but rather may have been introduced by the CML parameter estimation method.

While the aggregate results from a Monte Carlo simulation give an overall picture of simulation model performance, a single simulation of the prices from 1999 to 2003, such as that shown in Figure 5.7 below, better illustrates the simulated price process.

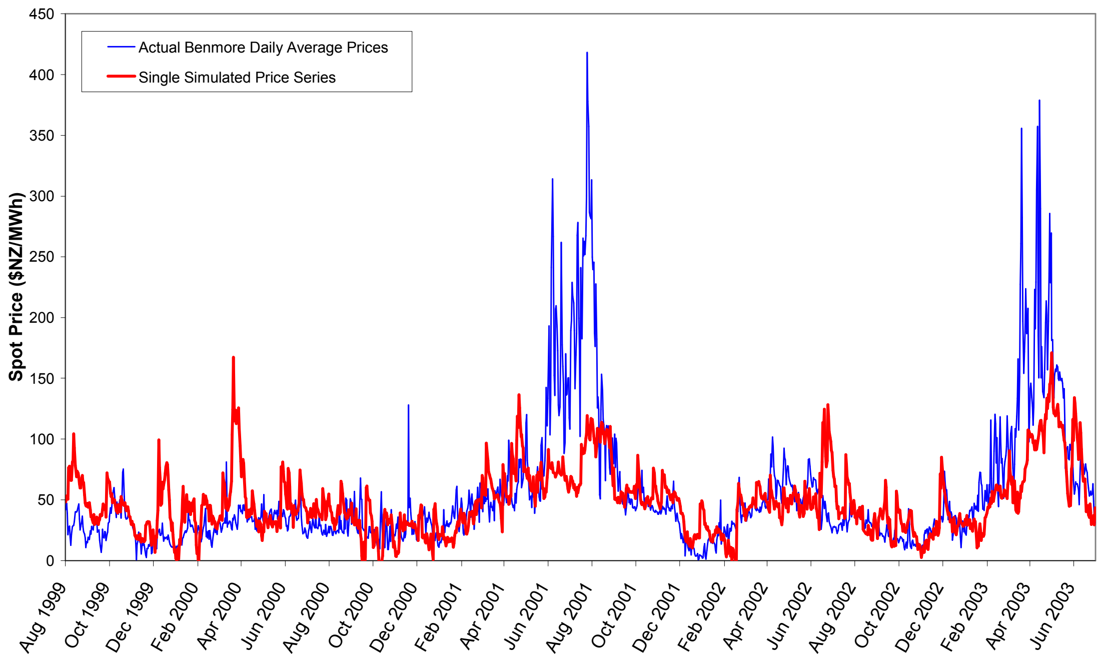

Figure 5.7: Daily average spot prices from the Benmore node, August 1999 – June 2003, and a single simulated spot price series over the same period using the MWV function price model

The price model is able to pick up the increase in the underlying price level in both 2001 and 2003; however several other observations can be made:

- Jumps in the simulated price series occur at random, and often at the wrong time of year.
- Simulated jumps are larger than they are in reality, and appear to occur much too frequently.
- Overall, the simulated prices appear to be far more volatile than they are in reality. This can best be observed in 2000 and 2002 in Figure 5.7, when there were large spikes and a great deal of volatility simulated despite there being a relatively large amount of water in storage. The actual prices may have increased and decreased slightly over time in 2002, but the level of inter-day volatility was much lower in reality than that simulated.

The reason behind these discrepancies is that nowhere in the price model is there any account taken of the fact that a key feature of the volatility in price time series is its dependence on supply-side factors. As explained earlier, and shown in Figure 5.3, it appears as though the inter-day volatility increases as the RSL decreases. Therefore incorporating some measure of the RSL into the stochastic component of the price model would seem to be a relevant extension to the model.

It may also be advantageous to remove all time-dependencies from the model. Currently the only change made to the EPV model has been the replacement of the original deterministic component with two seasonal functions for the MWV. The use of two MWV functions rather than one certainly improves the fit of the model, and appears to be reinforced by the observed data. However, having two seasons governed by the month of the year assumes that fundamental changes in behaviour occur both at the end of February and the start of September. While having one water value function would no doubt decrease the goodness-of-fit somewhat, it would remove this behavioural assumption, and enable the deterministic level of the price to be based entirely on the RSL and not on the time of year[5]. This will also produce a more parsimonious and intuitive model[6]. Therefore, the *j* subscripts are removed from equation [2b] in the model.

The other time-dependence that EPV impose on their model is the seasonal variance in the jump distribution parameters. Again, there is no practical reason for these parameters to change suddenly overnight, and these too appear as though they should be based

[5] But note that the calculation of the RSL itself is time dependent, so the MWV still depends on the time of year.

[6] Furthermore, recent NZEM price behaviour since June 2003 suggests that prices in the warm seasons of 1999-2003 did not reach high levels purely because there were no dry sequences of inflows in these periods. After the end of the sample period used in this research, New Zealand experienced another period of low inflows, but this time it occurred between September 2005 and January 2006, when inflows would normally have been increasing. This left reservoirs empty after the winter, at a time when they would normally be refilling in preparation for the following winter. Spot prices reached $140/MWh in November, climbed to $200/MWh in December and were still averaging around $90/MWh in mid-January when inflows began increasing above average levels.

entirely on the storage level. EPV justify varying these parameters seasonally due to jumps being more likely when either demand is high, or the amount of excess capacity is small. While load in the NZEM does vary seasonally, it is far more likely that a shortage of supply is more of a prerequisite for the occurrence of jumps. Therefore, as the MWV can be interpreted as a measure of excess capacity (the higher the MWV, the less excess capacity available today), it makes sense to test this exogenous variable as a driver of price volatility.

5.2.2 Incorporating the MWV into the stochastic component

Without complicating the stochastic component to any great extent, the elements that could incorporate the RSL are the unconditional variance parameter in the GARCH equation ω, the jump probability λ and the jump mean μ and variance σ^2. Due to the fact that the RSL can often be negative, the simplest method of incorporating the storage level into these processes would be to use the estimated MWV instead. Although it could be negative, it is practically very unlikely ever to be so. The simplest method of incorporating the MWV into the stochastic component is to make each of these parameters a linear function of the MWV. For example, the jump probability for each day can be calculated as the sum of a constant parameter and a parameter multiplied by that day's MWV; if that parameter is positive, as one would expect, then the greater the MWV the higher the probability of there being a jump in the price.

Including exogenous variables in GARCH processes for the conditional variance of asset prices is common (see Engle, 2002), and requires a minor adjustment to the likelihood function (detailed in Appendix B) with which the parameters in the price model are estimated. Adjusting the likelihood function to include the MWV as a driver of the jump process is similarly straightforward, and requires only straight substitutions for the adjusted parameters.

Following the replacement of the two MWV functions in the model with a single MWV function, the following four adjustments are made to the price model individually:

1. Unconditional variance$_t$ = $\omega + \omega' \times MWV_t$
2. Jump probability$_t$ = $\lambda + \lambda' \times MWV_t$
3. Jump mean$_t$ = $\mu + \mu' \times MWV_t$
4. Jump variance$_t$ = $\sigma^2 + \sigma^{2\prime} \times MWV_t$

The parameter estimates, estimated t-statistics and overall model fitting statistics for the base model with a single MWV function are shown in Appendix D. Note that both the log-likelihood and the SIC have decreased after the imposition of a single function for the MWV has been made, which confirms that the fit of the single-MWV function model is inferior to the model with two MWV functions[7].

Using the GAUSS package, each of the four adjustments listed above is made individually to the price model, to assess whether each adjustment improves the fit to the observed prices and whether each new estimated parameter is statistically significant. The estimated value of each new parameter, the estimated t-value of each parameter and the goodness-of-fit of each adjusted model (as measured by the log-likelihood and Schwartz Information Criterion) are displayed in Table 5.1. Recall that a greater log-likelihood and/or a lower SIC indicates a better fit.

[7] A model with constraints imposed on some of the parameters will never produce a better fit than a model without constraints; the model with the single MWV function imposes the constraint that the MWV function parameters must be the same in both periods.

Adjustment	**Parameter**	**Estimated Coefficient**	**Estimated t-value**	**Log-Likelihood**	**SIC**
Base Model				-5103	10301
Unconditional variance ω'				-5103	10308
	ω	1.8343	27.6675		
	ω'	0.0295	0.0155		
Jump probability λ'				-5099	10299
	λ	0	0		
	λ'	0.0024	3.2450		
Jump mean μ'				-5100	10301
	μ	-1.3897	-131.2010		
	μ'	0.4320	1.0834		
Jump variance σ^{2}'				-5101	10304
	σ^2	298.6015	811156.4		
	σ^{2}'	0.253935	0.623959		

Table 5.1: Estimated parameters, t-values log-likelihood values and SIC values from adjustments to price model

The adjustments to the model yield an interesting mix of results. Each of the adjustments increases the log-likelihood function, which is entirely as expected due to the fact that the fit of a model will always be improved if extra explanatory variables are added. As expected, the coefficients for all the parameters multiplied by the MWV are positive, which confirms that price volatility increases as the MWV increases. However, the fits to the model (as judged by the SIC) are only improved for the adjustments to the jump probability and jump mean size.

The estimated parameters λ and λ' for the jump probability clearly illustrate that the probability of a price jump is related to the MWV, therefore this change to the model is made permanent. The constant term λ is dropped from further modelling, however, due to its statistical insignificance.

Adding the ω' term to the volatility equation clearly does little to improve the fit of the model, and its estimated coefficient was statistically insignificant, therefore this adjustment was not made permanent. This result was not expected, but it illustrates that non-jump volatility is constant, and not related to the MWV.

The adjustments to the jump mean and jump variance both improve the fit of the model (as measured by the log-likelihood). However, because the estimated coefficients for the terms multiplied by the MWV are not significant (possibly due to both terms accounting for similar variation in the prices) further tests are required to determine whether the constant term or the MWV-multiplied term should be included in the final model in each case.

The original model (specified at the start of this chapter) includes just the constant terms, and yields an SIC of 10301. As shown in Table 5.2, when the constant μ term is replaced by $\mu' \times MWV_t$, the fit of the model is improved and the estimated coefficient of μ' is significant. The constant term μ is dropped from further modelling, due to its inclusion with the μ' term resulting in an inferior fit. However, the corresponding adjustment to the jump variance yields a worse fit, despite the statistical significance of the σ^{2}' parameter.

Adjustment	**Parameter**	**Estimated Coefficient**	**Estimated t-value**	**Log-Likelihood**	**SIC**
Jump mean	μ	13.4738	1312.2353	-5103	10301
	μ'	0.4138	101.5037	-5099	10292
Jump variance	σ^2	323.9103	1056360.3	-5103	10301
	σ^{2}'	10.1821	1263.2883	-5104	10302

Table 5.2: Estimated parameters, t-values log-likelihood values and SIC values from adjustments to price model

As a result of this testing, we can conclude that while the unconditional GARCH variance and the jump variance are constant, and unrelated to the MWV, both the probability of a jump and the mean jump size increase as the MWV increases.

Given that the adjustment of both the jump probability and jump mean size were successful, the final test is to assess whether or not adjusting them at the same time improves the fit of the model. The results of fitting this model are displayed in Appendix D, with the crucial statistics being an improved log-likelihood of -5096 and improved SIC of 10287.

As expected, introducing both the adjusted jump probability and the adjusted jump mean reduces the log-likelihood and improves the fit of the model to a greater extent than when they were each introduced individually, and both parameters' estimated coefficients are statistically significant. How the jump probability and mean jump size change with the RSL are shown in Figure 5.8 below, along with the previous constant probabilities and mean jump size. Note that the mean jump size and jump probability are below their original constant levels for RSL above approximately 400GWh, and are above their original constant levels for lower RSL.

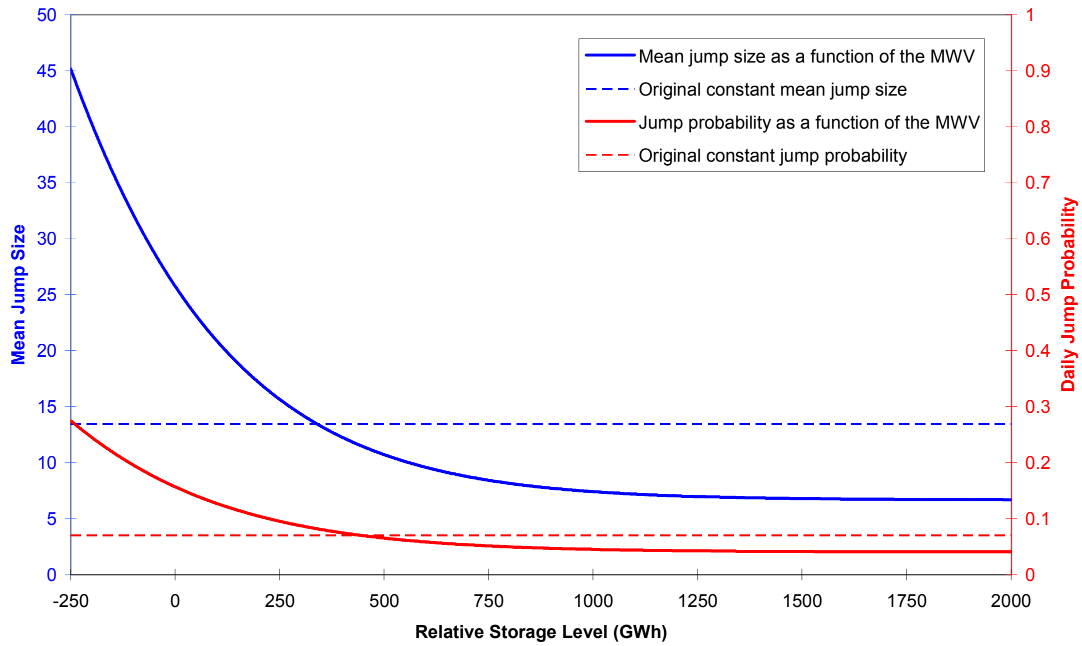

Figure 5.8: Comparison of jump probability and mean jump size from original and adjusted models

Running Monte Carlo simulations with the model including these adjustments to the jump probability and mean size shows how they have altered the performance of the model in terms of simulating time series over the length of the sample period. Figure 5.9 below, illustrating the simulation performance of the adjusted model, is the equivalent of Figure 5.6 for the original model. This figure shows that the median price in the simulation still follows the overall movements of the prices, apart from the two high-price periods. The 95% simulation bounds, however, are now narrower for the periods in which the MWV is low (e.g. 2000 and 2002), offering a reasonably precise simulation bound, while the bounds are wider in the periods in which the MWV is high. This is due to the fact that there is now less simulated volatility for periods when the RSL is high, and more when the RSL is lower.

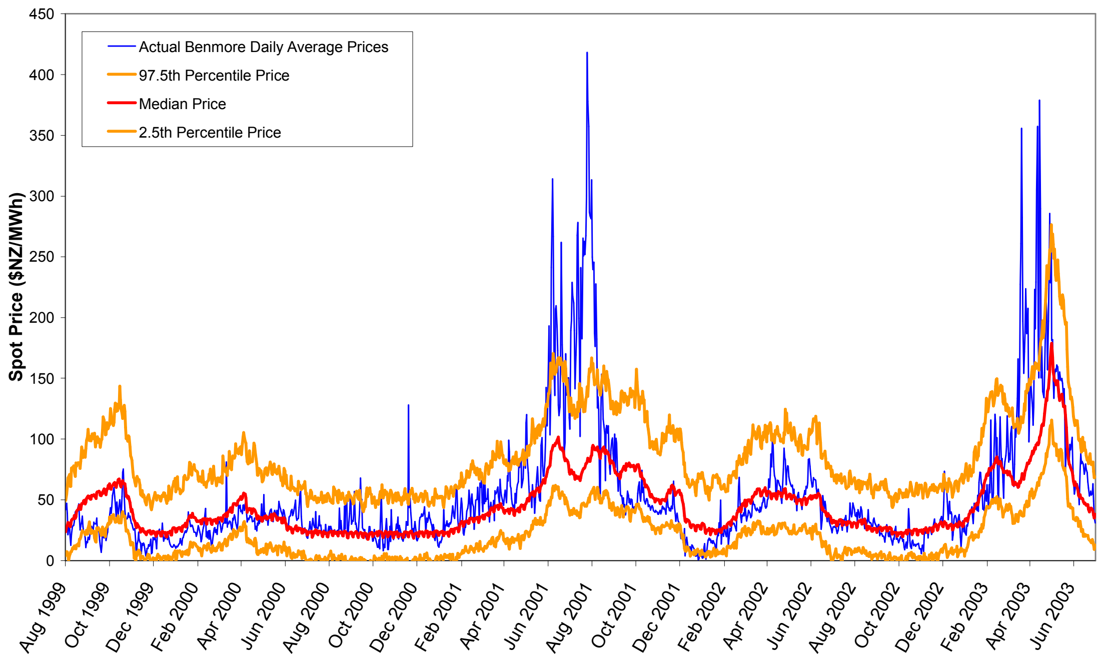

Figure 5.9: Daily average spot prices from the Benmore node, August 1999 – June 2003, and median, 97.5th percentile and 2.5th percentile simulated prices from running a Monte Carlo simulation over the same period using the MWV function price model with adjusted jump parameters

The charts of the simulation bounds before and after both the adjustments are combined in Figure 5.10. The green lines show the new simulation bounds and the orange lines are the original bounds. It is evident that the timing of the simulated volatility is now more appropriate. However, due to the apparent inability of the CML procedure to estimate parameters for the model which allow the price to increase as much as required in 2001 and 2003, the price is still underestimated in these periods and the observed prices lie outside the prediction bounds[8].

[8] Just why the CML procedure has this inability is unclear, however it suggests that any future research should make use of a different method of parameter estimation. It is worth noting again that the parameters fitted by hand in the previous chapter *were* able to capture the price increases in 2001 and 2003, however, when they were given as input to the CML procedure, they yielded a lower log-likelihood than the final parameters CML converged upon, and the search procedure moved away from that starting solution to the final solution provided in Table 4.. These parameters may have appeared by eye to

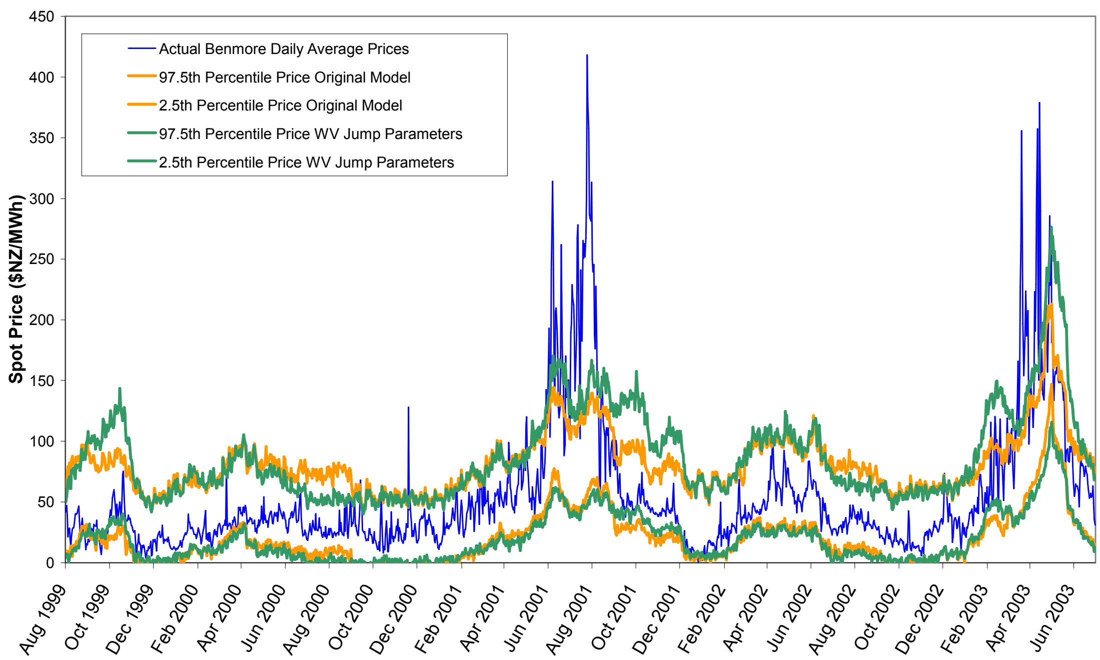

Figure 5.10: Daily average spot prices from the Benmore node, August 1999 – June 2003, and a comparison of the 97.5th percentile and 2.5th simulated percentile prices from running a Monte Carlo simulation over the same period using the original MWV function price model (with two MWV functions) and the single MWV function price model with adjusted MWV jump parameters

While little information can be gleaned from a single simulation of prices over the sample period, Figure 5.11 shows at least that the price model is able to simulate a reasonably accurate series of prices. The major benefit of the adjustments has been to ensure that the timing and size of the simulated volatility is more appropriate. Price spikes are now larger and more frequent when the MWV is high, which reflects the situation in reality.

---

produce a better fit to the price data, however, as explained in the previous chapter, they had a lower likelihood of being correct given the process specified for the price time series.

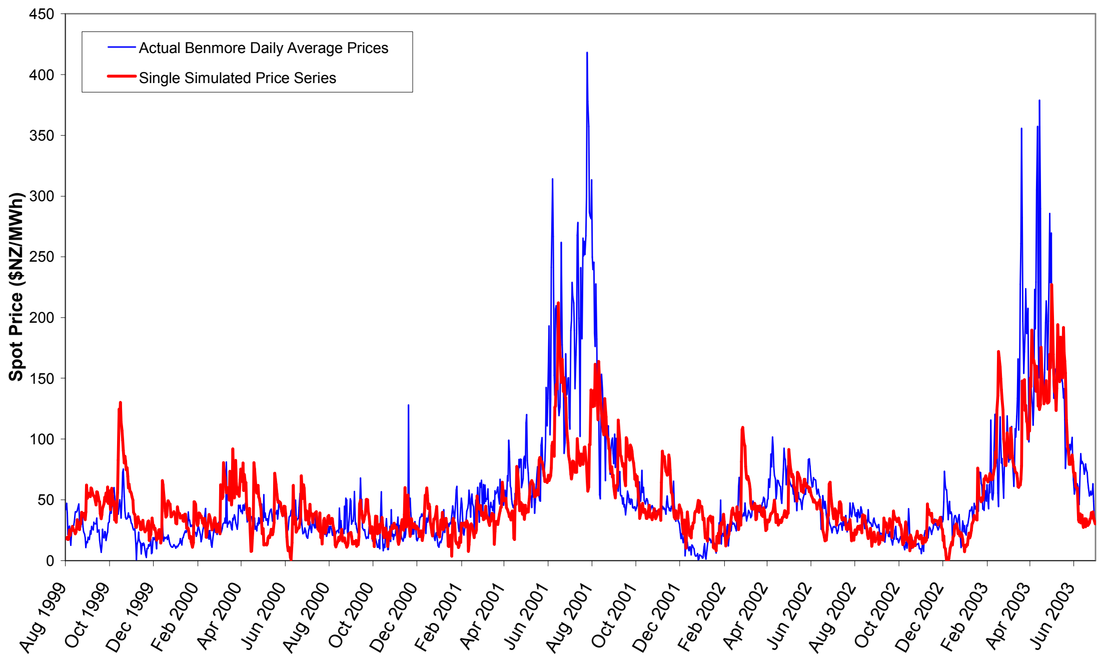

Figure 5.11: Daily average spot prices from the Benmore node, August 1999 – June 2003, and a single simulated spot price series over the same period using the model including the single MWV function and the adjusted MWV jump parameters

5.3 Conclusions

The incorporation of a linear trend in the lower envelope used for calculating the RSL showed that the lower envelope is slowly increasing over time. This suggests that reservoir managers may be becoming more conservative over time, and more recently have been assigning higher values to the same levels of storage. This helps to explain why prices remained at low levels in 1999 compared with 2001 and 2003, even though the storage level was relatively low.

However, while the model including a linear trend in the lower envelope produced an improved fit to the price data, the estimated parameter value for the trend was not statistically significant. Therefore the trended envelope is not included in the price model used later in this thesis. We note that even if the trend had been statistically significant between 1999 and 2003, there is no way of knowing whether or not the trend will continue into the future, and after another ten or so years it may be that the lower envelope reaches levels higher than the actual capacity of the Waitaki reservoirs, which in nonsensical.

In the previous chapter, the requirement to include the MWV as a driver of the level of spot prices was justified, and in this chapter the MWV has been included successfully as a driver of price volatility. Surprisingly, incorporating the MWV as a driver of the unconditional variance in the GARCH process and the variance of jump sizes did not improve the fitting power of the overall model. However, using the MWV to model the probability and average size of jumps was successful. This result is expected, due to the fact that price spikes have a higher likelihood of occurring, and are likely to be larger, when the amount of excess generating capacity (as measured by the MWV) is less.

For the rest of the NZEM price modelling in this thesis, the final price model used is the same as presented in the previous chapter, with the following adjustments:

- Only one function for the MWV is used, instead of two seasonal functions.
- The jump probability is a linear function of the MWV.
- The average jump size is a linear function of the MWV.

The final price model is therefore:

$$P_t = MWV_t + X_t \qquad [1a]$$

$$MWV_t = c + w\mathrm{e}^{x(y+0.01RSL_t)^z} + wkd.D_t^W \qquad [2c]$$

$$X_t = \begin{cases} \theta X_{t-1} + e_t & \textit{with probability } (1-\lambda'.MWV_t) \\ \theta X_{t-1} + e_t + \mu'.MWV_t + \sigma\varepsilon_{2t} & \textit{with probability } \lambda'.MWV_t \end{cases}$$

where $e_t = \sqrt{h_t}\,\varepsilon_{1t}$, and $\varepsilon_{1t}, \varepsilon_{2t} \sim \mathrm{N}(0,1)$. [5a]

and $h_t = \omega + \alpha\, e_{t-1}^2 + \beta h_{t-1}$ [4]

6

MODELLING NEW ZEALAND'S HYDRO RELEASES

6.1 Motivation

The model for NZEM spot prices presented in the previous two chapters uses a function of the Waitaki hydro system's storage level to calculate both the underlying price level and the degree and extent of price stochasticity. It has been shown that the fits to historic market prices produced by the model are an improvement on those produced when no storage information is included. However, while it performs well in a backcasting context, using historic storage levels, the model's shortcoming is that it cannot be used to forecast prices. This is because it would require forecasts of the storage level to do so, and, by itself, the model is unable to produce such forecasts. Producing forecasts of future storage levels, in order to forecast spot prices using the NZEM price model, is therefore the focus of the research presented in this chapter.

The reservoir storage level on any given day in the future is driven by three key variables: the storage level today, the flows into the reservoir from today until then, and the releases from the reservoir over the same period. The starting storage level is a single value, and the storage level on any given day in the past or present is readily available. However, forecasts of inflows and releases are not easily produced, and are themselves affected by the interaction of many other variables, which makes the process of forecasting both inflows and releases reasonably complex.

6.2 Release and inflow modelling for the NZEM

Modelling inflows is an established area of hydrology research, and inflow modelling in the New Zealand context is well-established, with a recent report by Harte and Thomson (2004) providing statistical models for inflows around the country. Their report gives references to other inflow models used in New Zealand and overseas. After considering these studies, inflow modelling has been ruled outside the scope of our research.

In contrast, release modelling using time series methods does not feature highly in the academic literature. Bottom-up or fundamental models are the major tools used for modelling changes in storage levels; however, these are much less transparent and conceptually harder to understand than top-down models. They generally require a complex multi-period optimisation tool for determining water values, as well as information on all the other generating costs and capacities available, in order to calculate both market-clearing prices and dispatch quantities. They also require assumptions regarding the strategies and behaviour of market participants. Moreover, as with other bottom-up models, these models fail to take into account the time series aspects of the outputs.

One such bottom-up model that has found application in the NZEM setting in recent years is the MinZone model, used by the New Zealand Electricity Commission and described on their website[1]. This model takes a starting storage level (say, the storage level at 1 April 2006) and, a sequence of historic inflows from 1 April to 31 March (from any previous year), and calculates the resulting storage trajectory under a certain set of assumptions. It then repeats the process with all the 75 or so historic inflow sequences from April to March on record (i.e. since around 1930) and, given those inflows, determines the proportion of those sequences that will lead to the New Zealand reservoirs running dry and the NZEM running out of generation. The basic aim of the model is to determine the risk of current storage levels (or hypothetical storage levels) leading to shortages in electricity generation, given a relatively large set of possible inflow sequences.

The MinZone model includes forecasts of load for the coming year, and uses a dynamic equilibrium bottom-up model (Ellsoft's *EMarket* model[2]) to determine water value contours, market-clearing prices and generator dispatch for each day in the coming year. It operates under the assumption that all non-hydro generation in the NZEM generates at maximum output levels, with only the residual load met by the hydro generation. In this respect, the storage policy employed by MinZone is unlikely to reflect reality. As a result, it is likely that when reservoirs are full (and water values low), modelled hydro generation would be much lower than would actually occur, and storage would therefore be kept at higher levels than usual. In reality, contract levels would most likely dictate that hydro generators operate at higher levels than MinZone calculates. At times of low inflows, modelled storage levels would therefore be artificially high; however, this is balanced by the fact that in MinZone hydro generators are forced to release, no matter what their storage situation, to avoid blackouts. Again, this is unlikely to occur in practice, with generators extremely unwilling to run their lakes dry and offering generation into the market at levels far higher than their water values in order to avoid being dispatched. Hence MinZone is likely to underestimate releases when reservoirs are full, and overestimate when reservoirs are empty.

---

[1] http://www.electricitycommission.govt.nz

[2] http://www.ellsoft.com/emarket.htm

Rather than make such assumptions regarding market behaviour, it makes sense instead to observe how the market actually operates in terms of water values, prices and releases, then fit and calibrate a model using that historic data. The estimated parameters of the model will therefore represent how the market behaves given certain storage situations and sequences of inflows, instead of having to make assumptions regarding dispatch priorities. Developing a top-down model using market data to represent the obvious relationships between storage levels, water values and releases appears a logical way to reflect that market behaviour.

6.3 Top-down release modelling

At the time the research in this thesis was undertaken, no other top-down release models existed in the academic literature to our knowledge; however, since then, Vehviläinen and Pyykkönen (2005) published a paper which simulates the operation of the NordPool market with aspects of both top-down and bottom-up modelling. Like the NZEM, the majority of power in the NordPool comes from hydro generation; hence their model is directly relevant to this research.

Vehviläinen and Pyykkönen model a market in which load is met by three types of generation: base-load supply (including unregulated hydro generation) regulated hydro generation, and condensing power generation[3]. They assume that "demand always exceeds baseload supply, and that the surplus demand is covered by regulated hydro-production and condensing power production". The spot price in their model therefore depends on the cost of those two different types of generation, and is set by whichever type supplies the marginal unit. The marginal cost of hydro generation (the water value, described below) does not vary with the level of generation, whereas the marginal cost of condensing power increases linearly with the amount of generation dispatched. For each

[3] A condensing power plant is similar to a standard steam turbine plant, except the steam passing through the turbine is collected and cooled inside a condenser and then returned to the boiler.

market clearance, the amount of condensing power dispatched depends on the price of hydro generation at the time. If condensing power is cheaper than hydro, then the amount dispatched increases until their marginal costs are equal, in which case hydro is dispatched until load is met or it reaches capacity, whichever occurs first. Any residual load is then met with condensing power.

The hydro water value (the marginal cost of hydro generation) is calculated as a function of the "normal" water value (a constant), the positive or negative change in the hydro balance from the last market clearance (the hydro balance is the sum of the reservoir level and the snow pack) and a penalty term that increases the closer the reservoir is to reaching its long-term minimum level. In contrast to the price models in this thesis, this minimum level is not time-dependent, despite the Norwegian storage being highly seasonal.

As with the model presented in previous chapters, the authors aggregate hydro storage into a single reservoir, and releases from that reservoir are calculated as a function of load (or residual load), the water value, and the price of condensing power. Releases will be greater if either the water value is lower or the load is higher, with all other factors held constant. Their model for load includes a fixed term for constant industry load, a temperature-dependent term and a noise term. They also model inflows explicitly by modelling temperature, precipitation, the rate at which the snow pack increases through precipitation, and the rate at which the snow pack melts. Each of the modelled parameters for the marginal costs, prices and loads are estimated from market data.

The model of Vehviläinen and Pyykkönen is highly detailed, yet still transparent and understandable. However, it has currently only been estimated with low-frequency data, and calculates one market-clearing price and set of dispatch amounts per month. As such, it captures neither the extreme price volatility that the model of this thesis does, nor the effects of extreme inflow sequences over a short period of time. In New Zealand, storage is much more limited than in Norway and, in a dry period, even one weekend's rain will have a pronounced effect on storage levels and spot prices. A model of at least weekly (but more suitably daily) releases is more appropriate for the NZEM for this reason. Secondly, their model again requires assumptions regarding the dispatch priorities of generating companies[4]. In the NZEM, each firm generally owns more than one type of generating technology; therefore it is unwise to make assumptions regarding the strategies they will employ. Finally, a key ingredient of their model is load, however at the time this research was undertaken we, like other researchers mentioned in this thesis, were unable to source a series of NZEM load[5], and modelling load is outside the scope of this research. In the course of this chapter, however, it will be shown that load may not actually be a necessary factor in determining hydro releases in the NZEM.

The aim in building a model for daily release is twofold. First, and most obvious, is the requirement to be able to model daily releases from New Zealand's hydro reservoirs, in order to forecast future storage levels more accurately. In doing this, we would like to incorporate as many relevant drivers of release as possible. Secondly, after hypothesising which drivers may be relevant, we would like to establish how much influence each of those drivers actually has on release, and therefore gain insights regarding the management of storage in New Zealand. One particular postulation in New Zealand today is that storage levels are kept at lower levels than in the pre-market era, because the incentive of power companies is to make profits rather than guarantee supply. Also, there are fears that high prices for electricity are likely to result in greater releases and lower lake levels as companies seek to maximise their profits, which will result in even higher prices. Analysing market data to assess how companies actually are operating will determine whether these fears have any substance.

Until this point, this thesis has focussed on the storage level from the Waitaki system in the South Island, and illustrated the link between this storage level and the Benmore prices. Generation assets on the Waitaki River are currently owned and managed by one

---

[4] While the dispatch priorities they assume are economically intuitive and consistent from the point of view of the market, they may not be internally consistent within individual firms.

[5] Recently, however, the New Zealand Electricity Commission has released its Centralised Dataset, which contains data on energy flows into and out of the grid.

SOE, Meridian Energy Limited. While the storage status of their reservoirs obviously has a major influence on the spot prices in the vicinity, their strategy for releasing water will also be highly dependent on factors such as their contract level, for which the relevant data is confidential. It is therefore not ideal to estimate a model for releases from a system of reservoirs managed by one company, without access to data on all the factors that influence the releases from that system.

For this reason, in order to compile a model that more accurately reflected the status of storage in the whole NZEM, we re-estimated the final price model from the previous chapter using aggregate storage data for the whole country[6] and prices from the Haywards node. As explained earlier, very rarely is the flow across the HVDC link between the North and South Island constrained, causing prices at Benmore and Haywards to be substantially different. However, when the flow is constrained, it is nearly always in the South-to-North direction due to a relative shortage of cheap generating capacity in the North, resulting in prices being higher in the North Island than the South. Prices at the northern end of the HVDC link therefore represent the current situation regarding the supply of generation to the majority of the load (which is in the North Island) more often than those at the southern end.

The price model referred to from this point on therefore models the final spot prices from the Haywards node using the national aggregate storage level. The estimated parameters for this model, along with comments on the comparison between these parameter estimates and those in the previous chapter, are given in Appendix E.

[6] We exclude the storage data from Lake Waikaremoana from this aggregate, due to the fact that its level has not been accurately recorded for some years.

6.4 The time series of NZEM releases

Because we were not provided with generation or release data explicitly, we were required to calculate releases from the series of daily aggregate storage and inflows for the whole of New Zealand. The formula for the total national release on day t, R_t, is:

$$R_t = Storage_{t-1} + Inflow_t - Storage_t$$

This formula states that today's release is the difference between the storage level at the start of today (plus the inflows for today) and the storage level at the end of the day. The owners of COMIT, the trading platform and information system of the NZEM, who provided the data, confirm that this formula gives a correct measure for R_t (Dean Yarrall, M-Co: personal communication, 2004).

It should be noted that R_t does not include non-storable release, such as tributary flows, that account for a proportion of the generation at some of New Zealand's hydro power plants. The diagram below in Figure 6.1 illustrates this further. Some precipitation in hydro systems flows through canals and rivers into storage reservoirs and may be stored for future release through the generating plant. The series of "inflows" used in this thesis records such "storable" flows into New Zealand's reservoirs. Other precipitation flows directly to (and through) the generating turbines, bypassing the reservoirs, and as such is "uncontrollable" in the way in which it can be used for generation. Tributary flows and controllable reservoir release then combine as generating release, and continue flowing down the river or canal below the generating plant[7]. For the purposes of this study, "release" is defined as controllable reservoir release, and "inflow" is defined as storable flows into a reservoir, as it is the reservoir storage level (and not total generation levels etc.) that we are aiming to model.

[7] The system becomes more complex if there is a linked series of reservoirs, however a one-reservoir representation is sufficient for the purposes of this thesis.

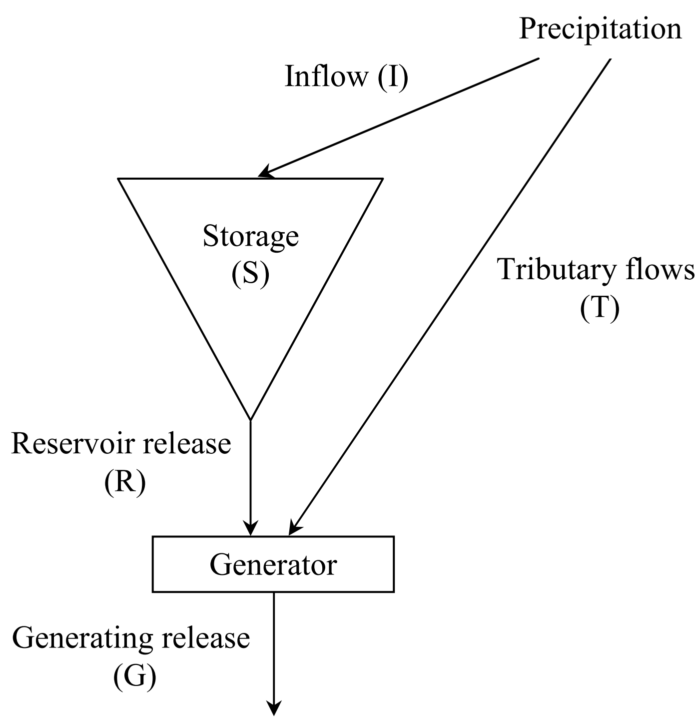

Figure 6.1: A representation of the relationships between inflows, storage, release and generation

The plotted time series of releases from April 1999 to June 2003 makes interesting viewing. As can be seen from Figure 6.2, the series is highly variable, and in order to stabilise that variation we took the natural log of the series for much of our further analysis. This is consistent with Harte and Thomson (2004), who model the natural log of inflows rather than the raw time series.

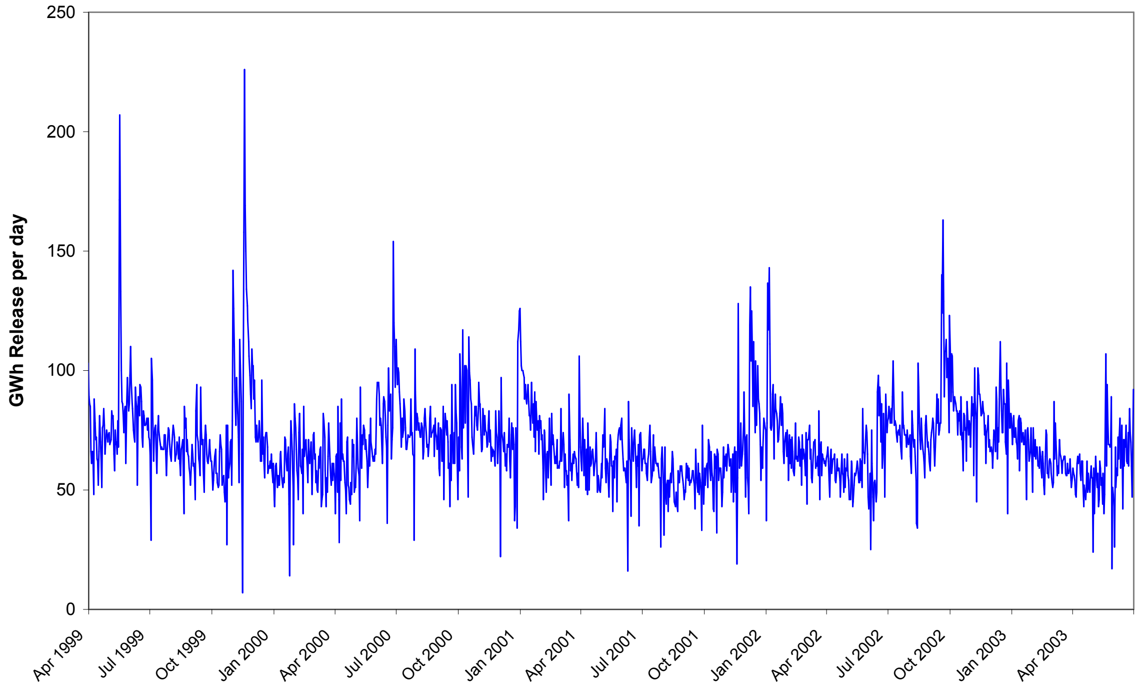

Figure 6.2: Daily aggregate releases for the NZEM, April 1999 – June 2003

Figure 6.3 below shows the average of daily inflow and daily release for each month over the length of the sample period, to give an indication of the positive relationship between the two series and the seasonal patterns.

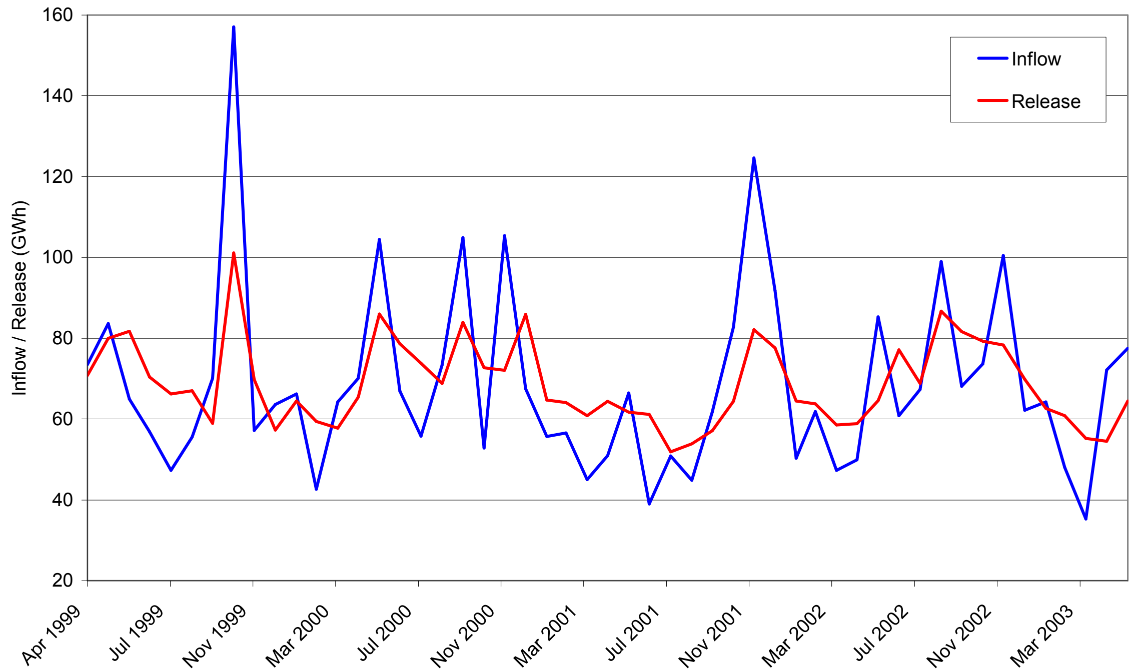

Figure 6.3: Monthly average aggregate inflows and releases for the NZEM, April 1999 – June 2003

The daily average releases per month appear much less volatile than the average inflows, which is to be expected as inflows are uncontrollable and releases controllable. There appears to be a strong positive correlation between the two series though[8], which suggests that when inflows are high, releases will also be high. Also, note from Figure 6.2 above that while there appears to be a seasonal aspect to the average monthly releases, with peak releases occurring in spring each year, this coincides with the high inflow periods due to annual snow melt. This is interesting, as it reflects the fact that New Zealand's storage capacity is limited. If this were not true, releases may be largest in the winter when load is at its peak to minimise the total annual cost of generation.

6.5 Drivers of hydro release

The first step in building a top-down model of release is to identify the important factors that may influence release. Each of these factors is listed below, along with a short

---

[8] This could be tested using cross-correlations, but this is not necessary at this stage of the analysis.

description of its relationship to release, and a simplified influence diagram[9], showing the relationships between all the factors, is provided in Figure 6.4. The selected decision variable for the influence diagram is a single day's release, under the assumption that the decision is made at the start of each day, given the information available at that time. While there are obvious feedback loops within the system, some of which are discussed below, for the purposes of the modelling in this study, these loops are ignored.

Current storage level

This is a known factor at the start of each day, and will likely have a large impact on the amount of water released. If storage is completely full, releases will potentially be large as any new inflows will have to be released to avoid spill. However, a generating company with an empty reservoir will be much more likely to want to conserve water, therefore releases will be less.

(Estimated) Marginal Water Value

This is expected to be a more significant driver of release than the absolute storage level discussed above. The marginal water value (MWV) reflects not only the absolute storage level but also the storage level with respect to expected seasonal variations in inflows and load. It also incorporates information regarding future inflows and spot prices, as well as flow requirements. As a result, and in contrast to the absolute storage level, the MWV reflects both the relative risk of running out of water and the risk of having to spill water. Therefore, if the MWV is high, water is relatively scarce and there is a higher risk of shortage later on in the year, and vice versa. The higher the risk of shortage, the lower release is likely to be; therefore, the MWV will likely have a major impact on release.

[9] The influence diagram (see Daellenbach, 1994) is a tool used in Decision Analysis, illustrating how different components in the "system of interest" influence (and are influenced by) the other components in the system. Uncontrollable events or inputs (i.e. outside the system of interest) are represented as clouds, system variables or calculations are represented as circles, and decisions to be made are represented as squares.

The MWV also reflects the impact that releases will have on the relative storage level (RSL) at the end of the day; if the MWV is high, any release will have a substantial impact on the resulting storage level (and hence the water value). If the MWV is low, more water can be released without increasing too much the risk of a water shortage in the coming weeks or months.

As we have discussed in previous chapters, we make the assumption that the water values used by power companies directly influence both the deterministic price level and the level and degree of price stochasticity. Since, in the context of our research, we do not have access to the water values actually used by power companies, we can only estimate what they may have been using market data. Therefore, for the remainder of this research, it is hypothesised that the estimated MWV, which is a function of the current RSL[10], is a direct driver of release.

Previous and current inflows

While the inflows from previous days are accounted for by the current storage level, they may also influence today's release level. For example, if there had been very high inflows into a reservoir on each day in the past week, it is possible that releases would have been large too, to ensure that the level of the reservoirs did not experience too great a change, both for environmental reasons and to reduce the risk of spill. However, in the case of the reservoir system in Figure 6.1, increased precipitation will increase both inflows (I) and tributary flows (T). In order to meet generation release (G) targets, reservoir releases (R) will then have to be reduced to balance the increase in T.

If there are several reservoirs on the same river, as is the case on the Waikato and Waitaki Rivers, the extent and exact nature of the relationship between previous inflows and current releases is unclear, regardless of whether or not there is any generation from

[10] The current RSL is also likely to be a driver of release. However, it was not included in this analysis as well as the MWV, which is a monotonically decreasing function of the RSL.

tributary flows. For example, inflows into any reservoir on such a river will be the combination of releases from the reservoir immediately above it and other flows that enter the river in between the two reservoirs. If heavy rain over the previous few days has both increased inflows into all reservoirs on the river over the previous few days and increased the likelihood of flooding on certain sections of (and in certain reservoirs on) the river, then releases from the reservoirs above the flooding sections will have to be limited so as to not increase that likelihood further. Releases from the reservoirs at risk of flooding will have to be increased, whether that is through spilling or generating. However if flows into the first reservoir on a river have been high, while those into the lower reservoirs have not, releases will have to increase from the first reservoir without the releases from the lower reservoirs necessarily being affected. Also, if a reservoir low on the river has had several days of high inflows and is at risk of having to spill water, it may be more advantageous for the generating company to reduce releases from reservoirs above it. Understandably, exactly how all these requirements will be represented in a single dataset combining data from many reservoirs is unclear.

Forecasted (future) inflows[11]

As with previous inflows, in the simple case with one reservoir it is likely that the more water is forecast to flow into the reservoir over the coming days and weeks, the more will be released. If the reservoir were nearly full and high inflows were expected over the next week, it is likely that current releases would be high to reduce the risk of having to spill water. If the RSL were low and high inflows were expected, there would not be the same incentive to increase releases; instead, the immediate risk of shortage would decrease, but releases would still be expected to increase with higher inflows. As with inflows on previous days, the likely relationship between forecasted inflows and current releases is much less clear for series of reservoirs on the same river.

---

[11] Instead of using actual historic inflow forecasts, we instead use as a proxy for those forecasts the inflows that actually occurred for each day in the coming week. These can be thought of as being perfect forecasts.

Level of contracts

Contract levels play a major role in determining optimal levels of output for a generating company. While a detailed discussion of the effects of contracting is outside the scope of this thesis, basically if a firm is contracted to supply electricity their incentives to generate electricity are vastly different compared with the situation if they were not contracted. The consequences for a non-contracted firm of reducing hydro generation through lack of water will be limited to lost revenue. However, a firm that is not able to fulfil its contractual obligations through its own generation will effectively be forced to purchase electricity produced by other generators from the spot market, in order to supply its customers. If the spot price is high, this can prove very costly – every day the firm is unable to fulfil its contracts will cost it money[12].

The overall significance of contract levels is that they play a huge part in determining releases. A firm that has a high level of contracts may release the same amount of water regardless of its relative storage level, as it does not have the same incentives to increase generation when water is plentiful (or reduce generation when water is scarce) as a firm that is not contracted. Unfortunately, however, despite their significance to the behaviour of the market, contract levels are not available publicly, and therefore we have to rule this driver outside the scope of our research.

Flow requirements

Restrictions on minimum and maximum river flows are strictly enforced in New Zealand. These have been introduced due to the concerns about the impact of flow variation on the environment, and also to ensure that the various users of the rivers (such as farmers, fishermen and other recreational river users) are able use the rivers as well. However, in our research we have aggregated all New Zealand storage into a single reservoir, with a

[12] See Batstone (2003) among others for a more detailed discussion of the effects of contracting on hydro generation levels.

single flow in and a single release out. The flow requirements are therefore not modelled explicitly, but could be expected to be reflected implicitly in the observed data.

System load

Theoretically, system load will influence the amount of release. Of course, aggregate system generation must equal load, and if NZEM generation were produced entirely by hydro then aggregate release (excluding spill) would match load exactly. However, only around 65% of generation in the NZEM is produced by hydro, and the amount of hydro generating capacity in reality depends much more on the amount of storage available than on the load. Therefore, when water is plentiful, the majority of generation comes from hydro, but when water is scarce a higher proportion of load will be met through thermal generation. As mentioned previously, due to not having a series of system load we are unable to determine just what influence load has on release.

Spot price

The spot price should have a major effect on release, although the nature of the relationship between the spot price and release is unclear. One would expect that when prices are high, generating companies could make higher profits by releasing more water. However, when prices are high, the water value is likely to be high too, and releases should be lower when the RSL is lower. With these two effects working against each other, what should be the case is that when the difference between the price and the water value is greater (or the ratio between the price and the water value is large), releases will be greater. So, for example, when water is plentiful but prices are high for some other reason, one would expect generating companies to maximise their profits by releasing as much water as possible.

6.5.1 The assumed sequence of events in release scheduling

Other system factors influence the spot price, such as the price and capacity of non-hydro generation (which also influence the MWV in reality) and transmission capacity. Importantly, there is also likely to be some feedback between the amount of release and the spot price. In a relatively small market such as the NZEM, a single firm can have some influence over the price, and therefore if a great deal of hydro generation is offered, it is likely that the spot price will be lower than if a small amount of generation was offered (depending, of course, on the price at which that generation is offered).In this research we do not model that feedback explicitly, and assume the following steps occur at the start of each day:

1. The MWV is calculated, based on the relative storage level at the beginning of the day.
2. The spot price is calculated as a function of the water value.
3. Given these pieces of information, reservoir managers calculate the amount of water to be released that day.

Therefore, we are not modelling the price and release feedback loop explicitly, assuming instead that the events occur sequentially rather than simultaneously. We also assume that there is no feedback between the release calculated in step 3 and the MWV for the beginning of the day from step 2, as we are assuming the MWV is calculated before the release occurs[13].

Given all the drivers listed and described above, the influence diagram for the decision of how much water to release from the reservoir in a single day is represented in Figure 6.4 below:

---

[13] In a future study it would be worthwhile to estimate firstly the amount of release on a given day, and then use a bottom-up model to assess the effect of that offered release amount on the spot price once the residual load is met with non-hydro generation.

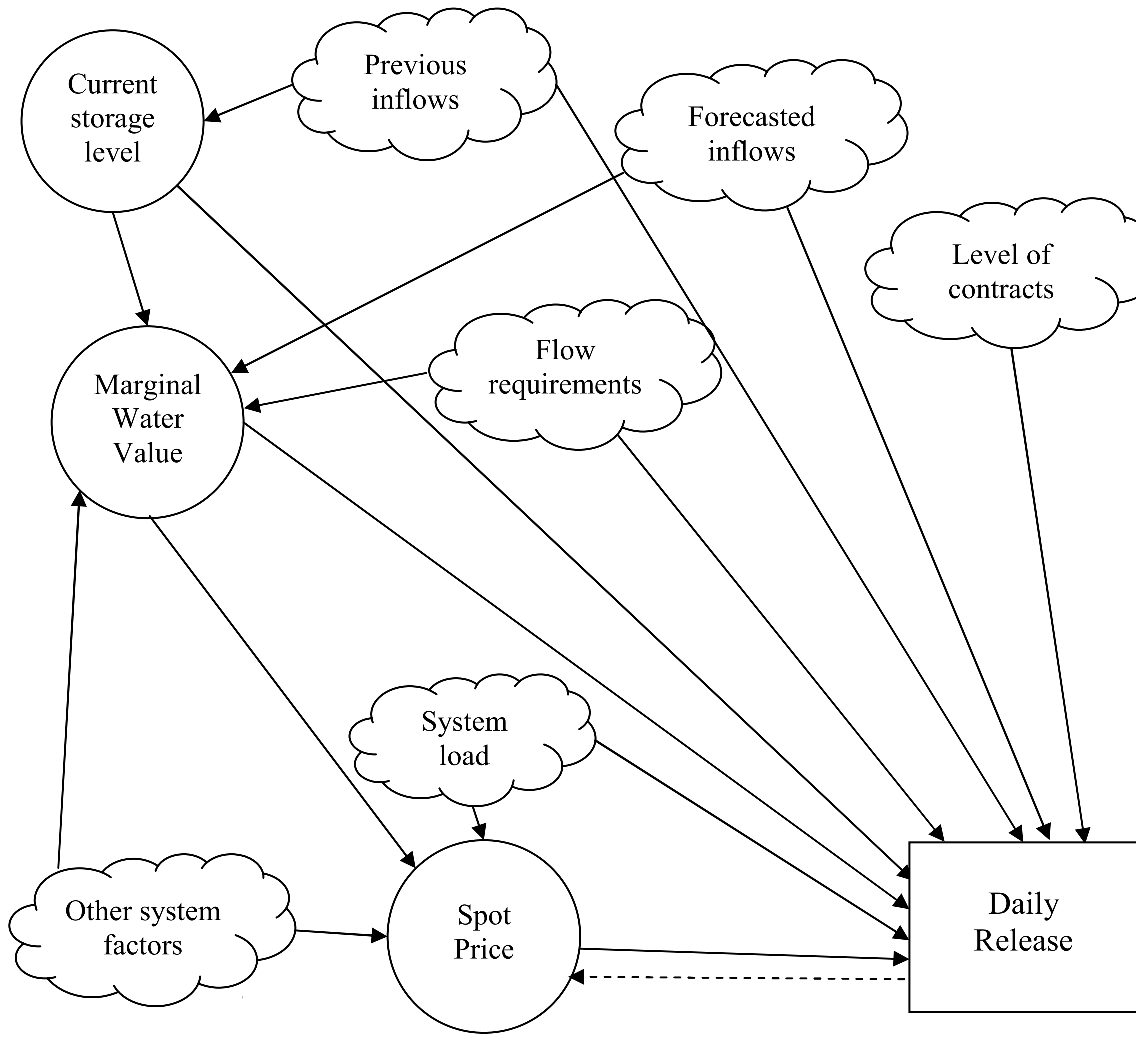

Figure 6.4: Diagram of factors influencing daily release from hydro reservoirs.

6.6 Initial investigations

The aim in estimating any top-down model is to develop a simple model that explains a large amount of the variation in the explanatory variable, and extend the model if necessary. The first step in assessing which of the drivers of release is significant enough to be included in a final model is to examine scatter plots containing data from individual drivers and release data. These establish which of the drivers appear to have significant relationships with release, and the apparent nature of those relationships determines the type of model that should be developed.

The first relationship illustrated is between the daily storage levels and the natural log of daily releases. The scatter plot of this relationship, shown in Figure 6.5, reveals a positive, roughly linear relationship: the greater the level of storage, the higher the level of releases is likely to be. However, the majority of the storage levels observed are between 2250 and 3250 GWh, and within this range there is a wide range of different levels of release.

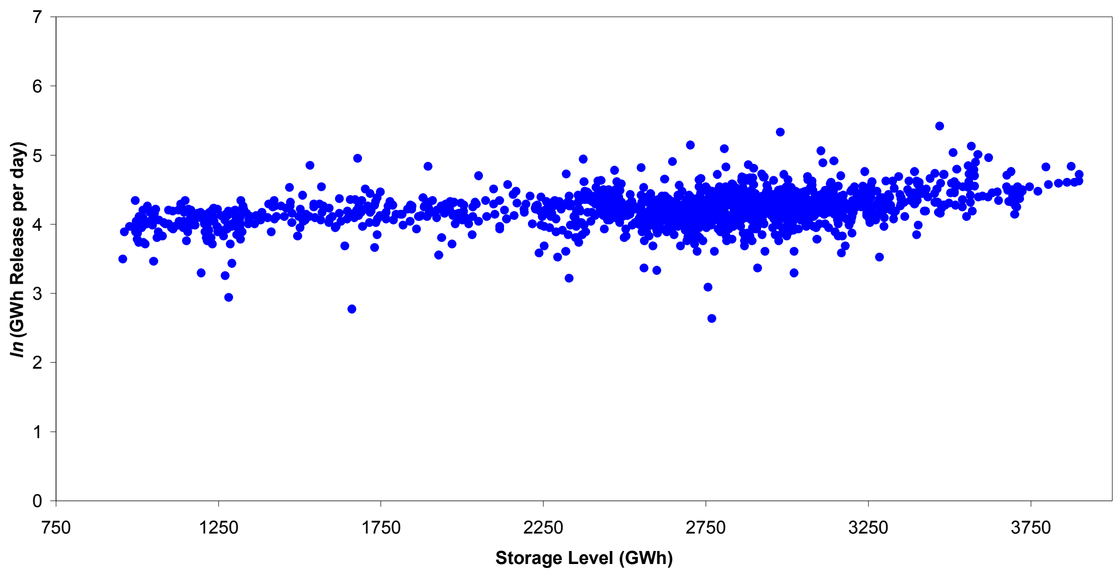

Figure 6.5: The national aggregate daily storage level versus the natural log of national aggregate daily releases, 1 April 1999 – 30 June 2003

Figure 6.6 shows the relationship between the estimated MWV and the natural log of release for each day in the sample. This relationship, as expected, is strongly negative, and appears linear as well. There is a wide range of estimated water values over the sample period, but the linear relationship appears to hold throughout the whole range.

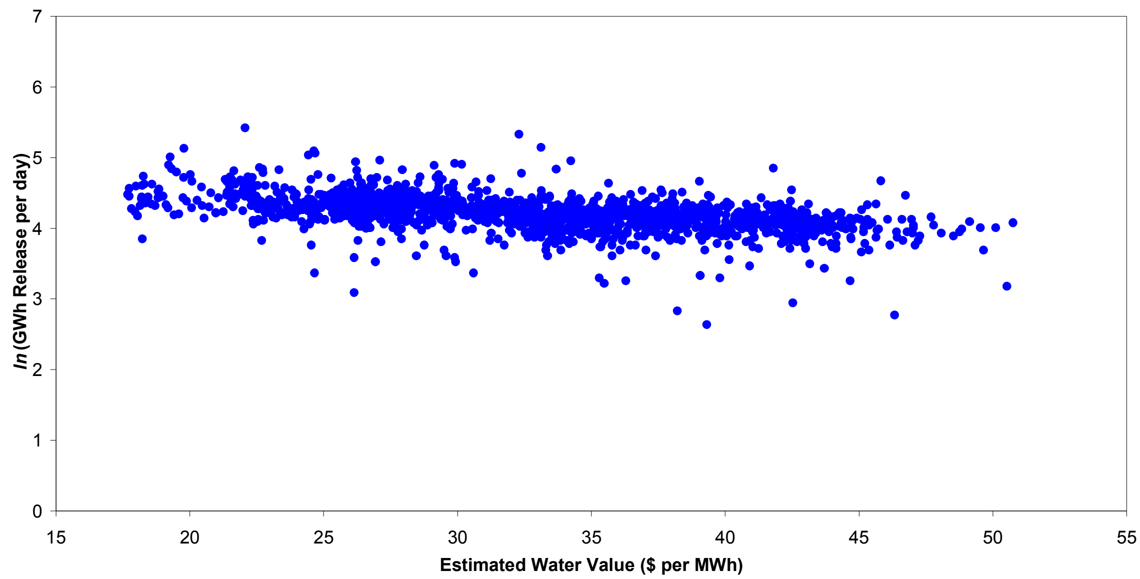

Figure 6.6: The estimated national aggregate marginal water value versus the natural log of national aggregate daily releases, 1 April 1999 – 30 June 2003

Releases are expected to be greater when inflows are greater, and this positive relationship is shown in Figure 6.7. This relationship also appears to hold throughout the range of inflows observed.

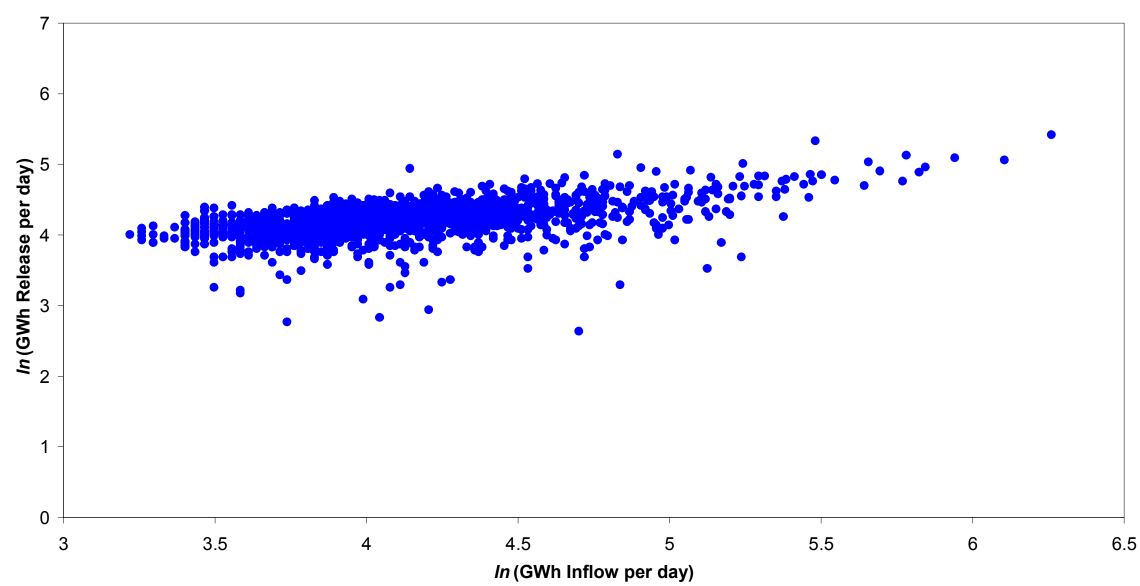

Figure 6.7: The natural log of national aggregate daily inflows versus the natural log of national aggregate daily releases, 1 April 1999 – 30 June 2003

One of the relationships expected to be strongest is the relationship between the spot price and the release, which is shown in Figure 6.8. It is hard to tell what (if any) relationships exist in this graph, as the majority of spot prices occurred between $0 and $100, and, if anything, the correlation between the two variables over this range of prices could be negative (see Figure 6.9). However, as prices increase above $100, the relationship, while still weak, appears positive.

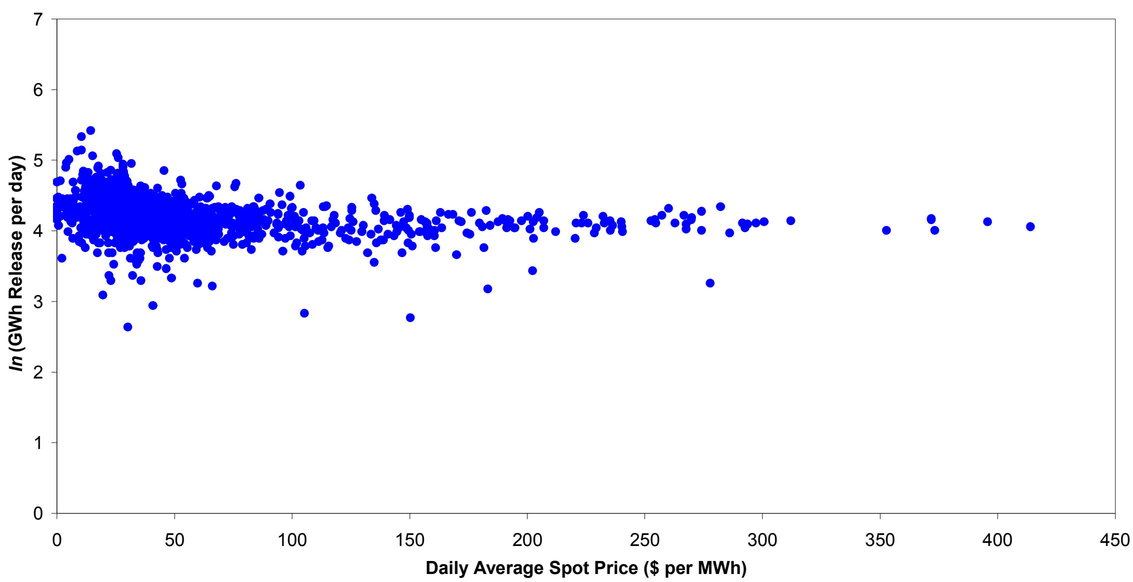

Figure 6.8: The daily average spot price at the Haywards node versus the natural log of national aggregate daily releases, 1 April 1999 – 30 June 2003

In order to show the relationship between the natural log of release and the spot price when the spot price is lower, Figure 6.8 has been reproduced in Figure 6.9 with the horizontal scale truncated. This reveals an apparent negative relationship between the spot price and release for lower levels of prices.

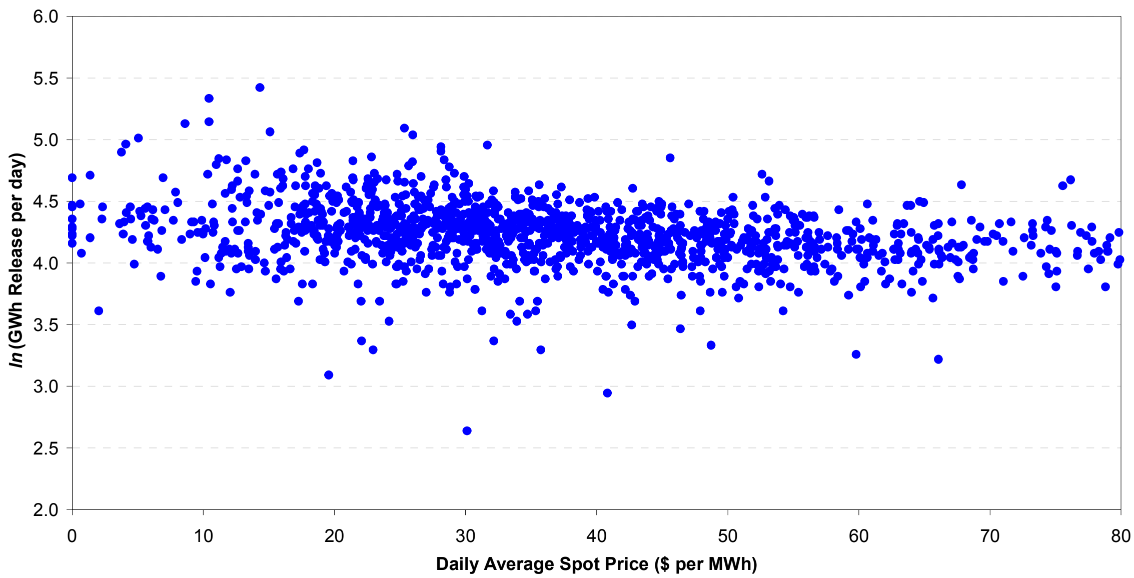

Figure 6.9: The daily average spot price at the Haywards node versus the natural log of national aggregate daily releases, 1 April 1999 – 30 June 2003. Horizontal (spot price) scale truncated to $0 to $80

As mentioned earlier, it was expected that the relationship between price and release would be strong and positive, especially on occasions when prices were significantly greater than water values. This scenario implies that prices are much greater than marginal costs, therefore the opportunity exists to increase profits by increasing generation. However, our investigations found little evidence to support this hypothesis, as illustrated in Figure 6.10. The red line shows the difference between the spot price and the estimated MWV at the time. Over the month shown in the graph (February 2003), this difference was always positive (i.e. the price was always greater than the MWV), and at times the spot price vastly exceeded the MWV. It was at times such as these that releases (the blue line) would be expected to increase significantly, however this does not appear to be the case with any consistency. Releases fluctuated only very slightly over the same period, while the price difference was highly variable. Neither does the counterargument, that decreasing releases in fact leads to an increase in the spot price, appear uniformly true. Thus we are forced to conclude that spot prices have much less influence over releases than we had initially thought.

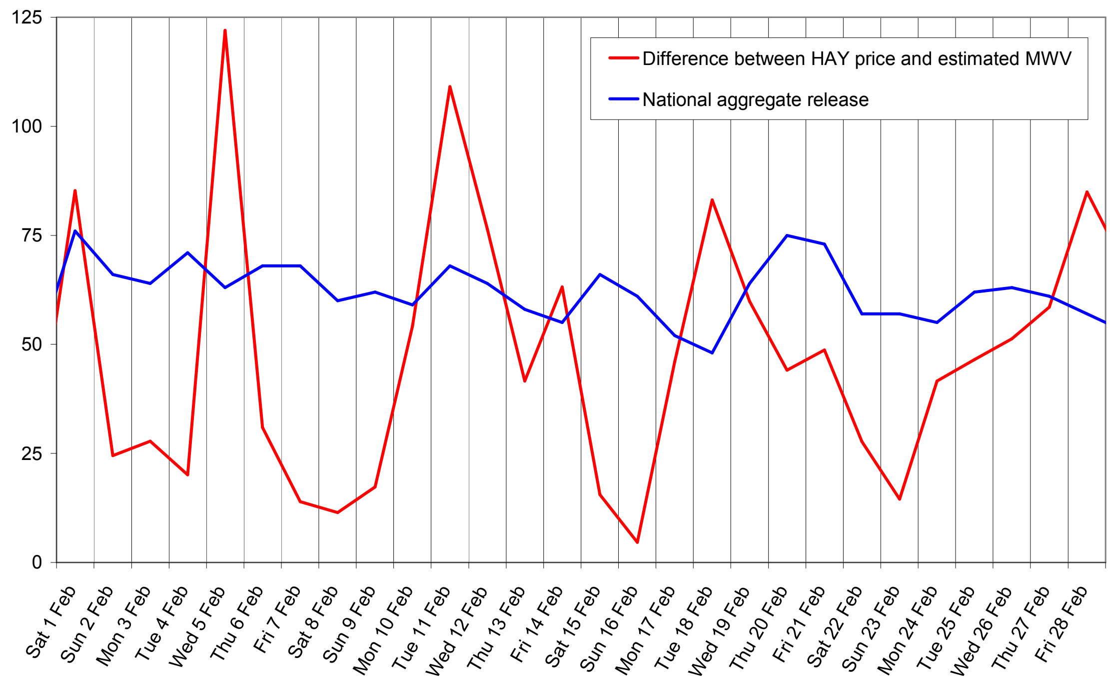

Figure 6.10: The difference between the daily average spot price at the Haywards node and the estimated national aggregate marginal water value (red line, in $NZ), and national aggregate daily releases (blue line, in GWh), February 2003

Due to the apparent linear relationships between the key drivers and release, as an initial step in the analysis a simple multiple regression was compiled of the natural log of release on all of the drivers we had available. Those drivers whose estimated coefficients were statistically significant would then be included in more advanced models. The complete set of input variables included in the multiple regression model for release on day *t* (from 1 April 1999 to 11 June 2003[14]) is shown in Table 6.1 below, as well as whether or not the estimated coefficient of each driver was statistically significant at the 5% level.

[14] The last 30 days of our available data, from 12 June – 11 July 2003, was withheld from the estimation in order to examine the forecasting performance of our final model.

Series / driver	**Statistically significant**	**Not statistically significant**
Constant	✓	
The natural log of the inflows prior to day *t*	Day *t*-2 Day *t*-1	Day *t*-7 Day *t*-6 Day *t*-5 Day *t*-4 Day *t*-3
The natural log of the inflow on day *t*	✓	
The natural log of the inflows subsequent to day *t*	Day *t*+1 Day *t*+2 Day *t*+3	Day *t*+4 Day *t*+5 Day *t*+6 Day *t*+7
Storage level on day *t*		✓
Price on day *t*	✓	
Estimated Marginal Water Value on day *t* (from the previously estimated Price Model)	✓	
Ratio of the price to estimated Marginal Water Value		✓
Series of dummy variables for each day of the week (i.e. for the Monday series, if day *t* is a Monday then 1, otherwise 0)	Saturday Sunday	Monday Tuesday Wednesday Thursday Friday

Table 6.1: List of significant and not significant drivers in the regression on the natural log of release

The results of this regression are shown in Appendix F. All the drivers expected to be significant proved to be, with the exception of the storage level and the ratio of the price to the water value. It is likely that the storage level is excluded due to the fact that the MWV includes all the relevant information regarding the storage level, and more. Also, including both the MWV and the price level as regressors implicitly includes the difference between these two variables as well; therefore, little extra information could be gained by including the ratio of the two.

When the natural log of release is regressed on all the drivers, the regression model has an adjusted r^2 value of around 56%, whereas when the level series of releases is regressed (and the level series of inflows are used as explanatory variables) the adjusted r^2 is around 70%. One interpretation of this result is that many of the extreme observations in the release series can be accounted for by corresponding extreme inflow observations. When the natural log of releases and inflows are used instead, the explanatory power of the inflow series is decreased.

As a crosscheck on the variable selection procedure, a stepwise regression[15] was run on the same set of data (see the results in Appendix F), with a tolerance level for selection of 5% significance. This method selected exactly the same list of variables as listed in Table 6.1, but it is interesting to note the order in which the variables were selected in the stepwise regression, and the proportion of the variance in releases for which each accounts. The variable chosen first was the estimated MWV, accounting for 26% of the variance, followed by the inflow on day t (a further 12%), inflow on day t+2 (10%), inflow on day t-1 (1%), then the Sunday and Saturday dummy variables (1% each). The other variables, while still significant, each add less than 1% to the overall regression r^2 value. Interpretation of each coefficient follows later in the chapter (once the final model has been presented); however, note at this stage that while the price is a significant variable, it is not one of the first ones chosen in the regression.

[15] Stepwise regression is a method for selecting significant explanatory variables from a group of potential explanatory variables, some of which may not be significant.

6.7 Formal modelling

With several lagged observations of one of the explanatory variables (the natural log of inflows) included, the regression model from Section 6.6 can be classed as a Dynamic Regression (DR) model[16]. Following the methodology proposed by Pankratz (1991) for developing DR models with more than one explanatory variable, the variables that were not significant were removed and the model was re-estimated in SAS using maximum likelihood estimation[17].

With any regression-type modelling, if one or more relevant explanatory variables are not included explicitly in the model then the effects that they have on the dependant variable will appear implicitly as "noise". Due to the fact that several likely important drivers of release, such as load and the contract levels, are unable to be included in this model, it is highly likely that the residuals of the model will exhibit patterns that need to be accounted for by an error process. Examination of the sample autocorrelation and partial autocorrelation coefficients (plotted in Appendix F) reveals some serial correlation in the residuals, which are fit with an ARIMA[18] process according to the DR-modelling methodology proposed by Makridakas et al. (1998). As the results in Appendix F show,

---

[16] As mentioned in the literature review in Chapter 2, DR models have found application in modelling spot prices. For example, Nogales et al. (2002) model spot prices as a DR and include current and past observations of load as explanatory variables.

[17] See Appendix B for an explanation of maximum likelihood estimation.

[18] ARIMA (Autoregressive Integrated Moving Average) models were formalised by Box and Jenkins (1976, as cited in Makridakis et al, 1998, p. 312). These models include terms to account for multi-period autoregression in both the dependent variable and the model residuals. Some series must be differenced one or more times in order to ensure stationarity. Standard notation for ARIMA models is ARIMA(p,d,q), where p is the order of autogression in the dependent variable, q is the order of autoregression in the residuals of the model, and d is the order of integration of the series (or the number of times the time series must be differenced before it becomes stationary). An ARIMA(p,0,q) or ARMA(p,q) model is fitted to series that do not need any order of differencing.

an ARMA(3,1) process for the residuals of the DR model minimises the chosen fitting criteria[19] and leaves no serial correlation or other patterns in the residuals of the model.

Our final dynamic regression model for release on day t is as follows:

$$
\begin{aligned}
ln(\text{Release}_t) \quad = \quad & 4.3893 \\
& - \ 0.0232 \quad \text{Estimated water value}_t \\
& + \ 0.0006 \quad \text{Daily average spot price}_t \\
& + \ 0.0542 \quad ln(\text{Inflow}_{t-2}) \\
& - \ 0.1834 \quad ln(\text{Inflow}_{t-1}) \\
& + \ 0.4029 \quad ln(\text{Inflow}_{t}) \\
& + \ 0.1830 \quad ln(\text{Inflow}_{t+1}) \\
& - \ 0.4427 \quad ln(\text{Inflow}_{t+2}) \\
& + \ 0.1248 \quad ln(\text{Inflow}_{t+3}) \\
& - \ 0.0785 \quad \text{Saturday dummy variable}_t \\
& - \ 0.0862 \quad \text{Sunday dummy variable}_t \\
& + N_t
\end{aligned}
$$

$$
\begin{aligned}
\text{Where} \quad N_t \quad = \quad & e_t \\
& + 1.0123\, N_{t-1} + 0.1115\, N_{t-2} - 0.1442\, N_{t-3} \\
& - 0.8956\, e_{t-1}
\end{aligned}
$$

$$
\text{and} \qquad e_t \sim \mathrm{N}(0,\ 0.0235)
$$

As shown in Appendix F, each parameter estimate is statistically significant at the 1% level. The signs of most of the coefficients are as expected, with the exception of those of the inflow series. For example, it is likely that as the water value increases, releases decrease, but the higher the spot price, the higher the release. Also, releases are likely to be less during the weekend, due to the fact that the load is less, although one might have

[19] The fitting criteria chosen were the Schwartz Information Criterion and the Akaike Information Criterion, which are both functions of the log-likelihood. As evidenced in Appendix F, both of these criteria were minimized by including an ARMA(3,1) process for the residuals.

expected this effect to have been captured in the relationship with the price, as prices are lower in the weekend than during the week. The estimated price coefficient was obviously too small to capture that effect fully.

The hypothesis presented in Chapter 4 of this thesis is that the deterministic level of the NZEM spot price is determined solely by the water value. Every day, a stochastic price component (which may be positive or negative) is added to that deterministic price level to form the spot price. In the model for release, the coefficient for the estimated MWV accounts for both the MWV by itself and the MWV as a component of the price. Therefore, since this estimated coefficient is negative, an increase in the price due solely to an increase in the MWV will decrease releases, as expected, since the marginal cost of hydro generation has increased. However, an increase in the price above the MWV, due to a price spike or some other aspect of stochasticity, will increase releases as generating companies seek to make greater profits. In summary, if the price is substantially greater than the MWV then releases will be slightly (though not much) greater. A dollar increase in the MWV will decrease releases by much more than a [non-storage induced] dollar increase in the price level will increase releases.

Whether or not the estimated coefficients of the individual inflow series have any meaningful interpretation is debatable. The expected result would have been for the coefficients for the inflows before day t each to have had the same sign, and those after day t each to have had the same sign; however this is not the case. As noted earlier, the first inflow variable selected in the stepwise regression was the inflow on day t, and the estimated coefficient for this series is positive, as expected and as shown in Figure 6.7. This shows that the more water flowing into the reservoirs today, the more will be released. Also, the sum of all the estimated inflow coefficients is greater than zero, so the net effect of a higher level of inflows over (say) a whole week or month is that releases will increase over that week or month, as expected and as shown in Figure 6.3.

The lack of clarity regarding the effect of previous and forecasted inflows on release levels is not surprising, given the background to each relationship provided in Section

6.5. In New Zealand, most reservoirs have very limited storage, and if higher inflows are forecast then the reservoirs have to be emptied somewhat in order for the inflows to be accommodated. However, for other reservoirs, if high inflows have been experienced in the previous few days (and are forecasted to continue) then releases must decrease to reduce the risk of flooding downstream.

Most rivers have lower limits on downstream flows, therefore releases will have to be of a sufficient level so as not to breach those limits, regardless of how low the inflows over previous days might have been. Still other rivers (such as the Waiau River above, and those rivers below Lake Manapouri) have restrictions on water clarity – if a river is in flood after heavy rain, then more water must be released from the reservoirs upstream in order to improve the downstream clarity. With all these factors influencing the relationships between inflows and releases, the mixture of estimated coefficients is not unexpected, hence we do not attempt to interpret each coefficient further.

The major interpretation of all the estimated coefficients in the model for release is that they should be considered as representing 'market' behaviour. They represent how the NZEM actually behaved, given the inflows and storage levels experienced between 1999 and 2003. Therefore, while some caution should be exercised when applying models to different data sets, the release model can be applied to a different set of data to examine how the market might behave in different situations.

6.8 Forecasting performance

Testing the forecasting performance of the release model over the 30-day holdout sample from 12 June – 11 July 2003 reveals that the release model performs exceedingly well. Figure 6.11 below shows the actual releases that occurred over that time period (blue line) along with the forecasted releases (thick red line) and a 95% prediction interval (dashed red lines). Only one day out of the 30 was outside the prediction interval, with the other 29 actual daily releases well inside. The forecasts track the path of the actual releases very well.

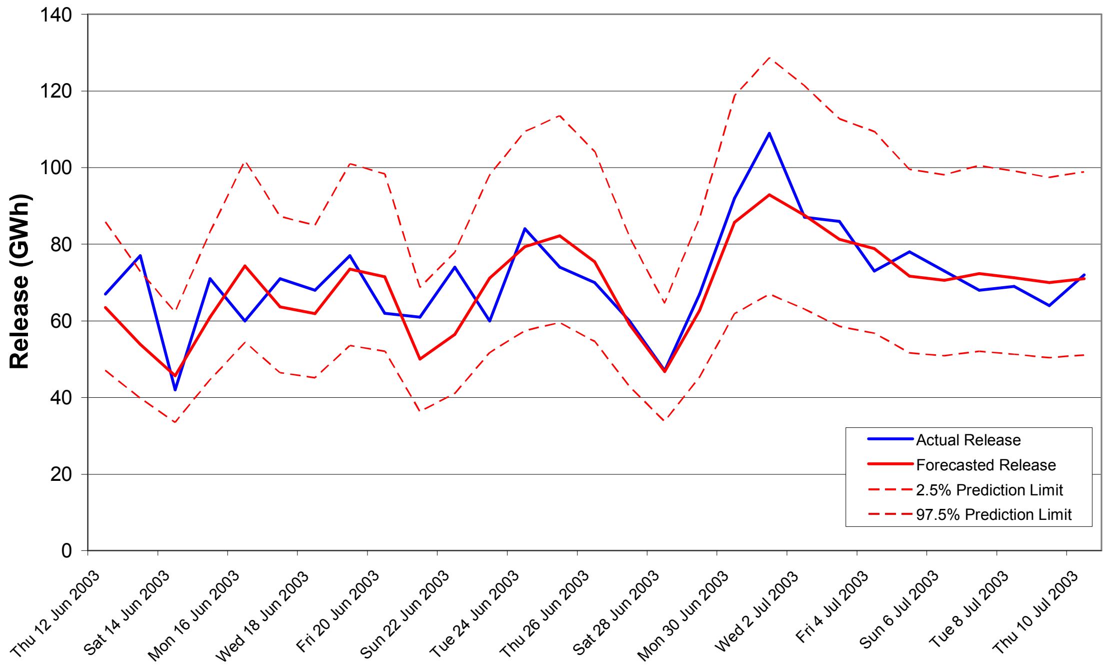

Figure 6.11: Forecasts and prediction intervals for the holdout sample using the model for release (red lines) and actual aggregate releases (blue line), 12 June 2003 – 10 July 2003

6.9 A combined price and release simulation model

In order to be used for forecasting releases, the DR release model presented in Section 6.7 requires only forecasts of the estimated MWV, the spot price, and a series of inflows. As mentioned earlier in this chapter, inflow modelling is well-established, and there are many series of synthetic and historic inflows available. When combined with a synthetic or historic inflow sequence, the release model can be used in conjunction with the NZEM spot price model presented in the previous chapters of this thesis to form a simulation model of storage levels and spot prices over any period of time. The only inputs required in the simulation are a starting storage level and a series of inflows. The steps in the simulation are as follows.

Given a starting storage level S_0 and an inflow series I, for each day t, starting from t=1:

1. Calculate the MWV_t from S_{t-1} using the relative storage level methodology.
2. Calculate the spot price, P_t, as a function of the MWV_t using the NZEM price model.
3. Calculate the release, R_t, as a function of the MWV_t, P_t and I_{t-2} to I_{t+3} using the DR release model.
4. Calculate $S_t = S_{t-1} + I_t - R_t$.

Using these steps we can backcast over the sample period from 1 April 1999 – 30 June 2003 to assess the combined simulation model's performance in forecasting storage levels over the period from which the parameters were estimated. The forecasted storage trajectories can then be used for forecasting prices, using the models presented in previous chapters.

Figure 6.12 below shows the actual storage levels from April 1999 to June 2003, as well as the median simulated storage levels and 95% simulation limits of the storage levels (i.e. 95% of simulated storage levels are within these levels). As can be seen, the simulated storage trajectory (red line) tracks the actual trajectory (blue line) very well over the majority of the sample period, with the exception of four continuous months in late 2000. During these months, actual releases were much less than simulated releases, which led the model to underestimate storage levels for a number of days. However, releases were vastly underestimated in January 2001[20], which led to the two trajectories converging again. 92.6% of the actual storage trajectory lies within the 95% simulation limits (dashed green lines), and, excluding those four months, only one day's storage level is outside the limits.

[20] This is discussed further in the following chapter, which analyses release behaviour in this period in more detail

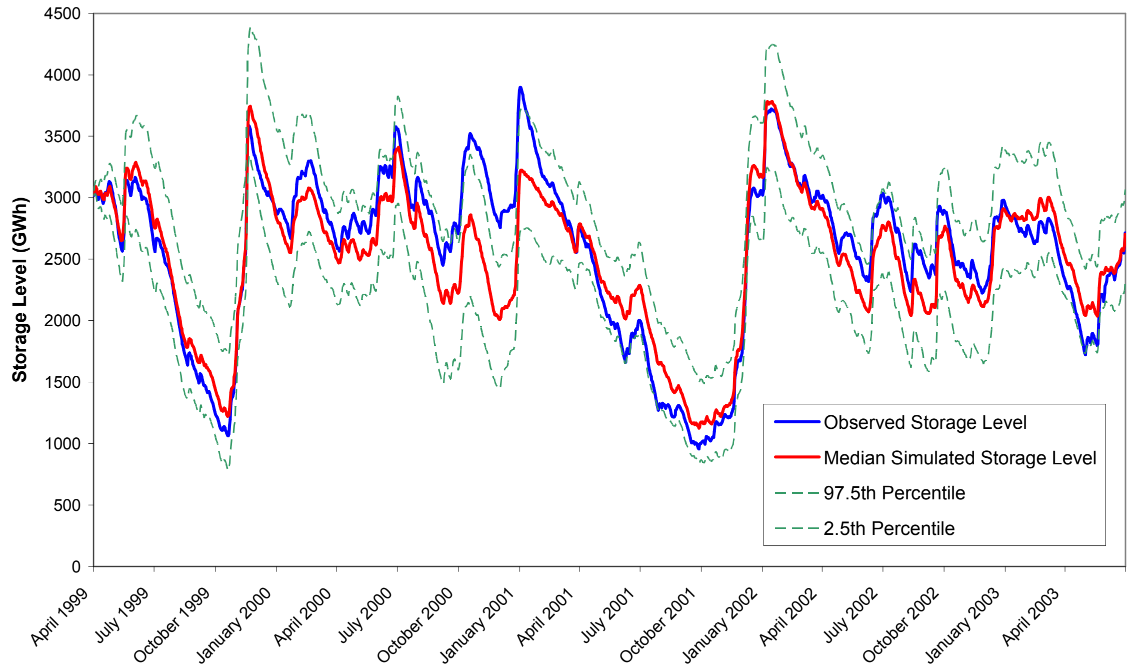

Figure 6.12: Actual aggregate NZEM storage trajectory (blue line), median simulated storage trajectory using simulation model (bold red line) and 95% simulation limits (dashed green lines), April 1999 – June 2003

Simulating storage trajectories as in Figure 6.12 allows simulated relative storage levels also to be calculated, and it is from these that water values and prices can be estimated.

As a further form of model and data validation, Figure 6.13 shows the average releases for each day of the week over the sample period, both real and simulated. The lower releases on Saturdays and Sundays are clearly evident.

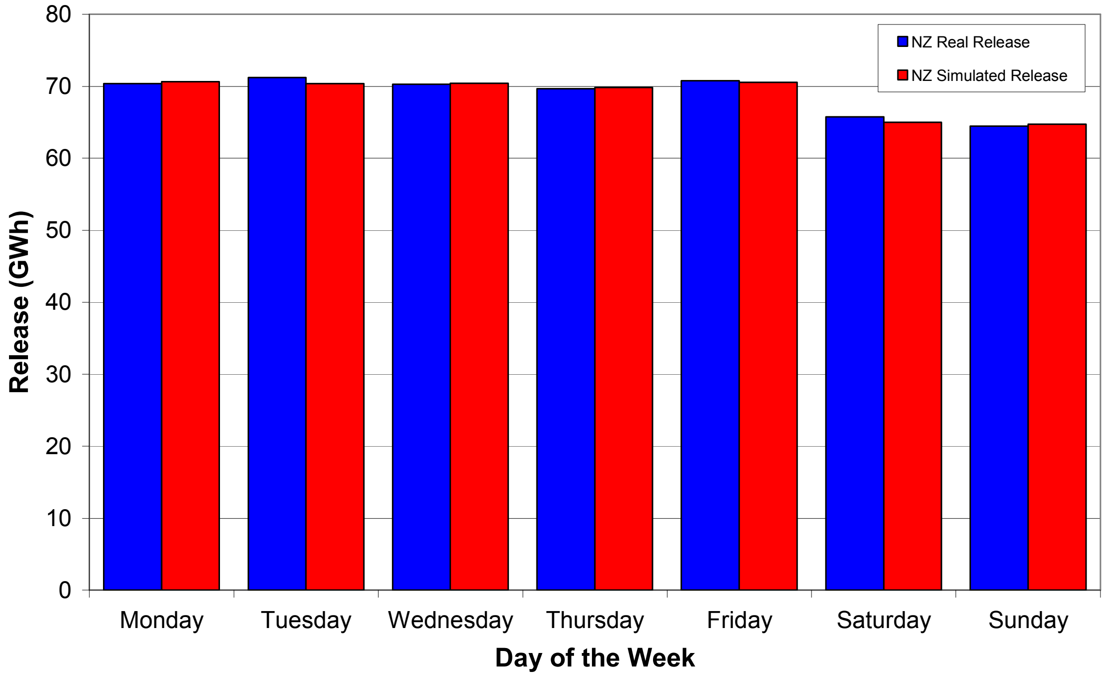

Figure 6.13: Average real releases per day and average simulated releases per day, April 1999 - June 2003

6.10 Conclusions

This study of release behaviour in a market context has resulted in some important insights into the way the NZEM operates. Most important of these is the fact that high spot prices do not necessarily encourage generating companies to release a large amount more water than they would have normally. On the contrary, the lower the RSL appears to be (and the higher the MWV), the less will be released, as generating companies act conservatively to avoid running out of water. Therefore, releases are driven to a much greater extent by the relative storage level and the inflows than by the spot price.

The effect of the spot price on release is so small that it could even be concluded that the dynamics of the spot market have very little influence on how the generating companies manage their reservoirs. However, this is not entirely true. One component missing from this model is the contract level of the hydro generating companies. In New Zealand, each of the major generating companies is vertically integrated to a degree, presenting each with a level of implicit retail contracts. As shown by Batstone (2003), risk-averse generating companies will generate to these contract levels, regardless of the spot price. Due to the dynamics of the spot market, a company has a limited incentive to deviate from this contract level. Assuming they have market power, if they generate more power, the spot price will decrease and each MW they produce will earn less revenue[21]; if they generate less, the spot price will increase and they will have to purchase power off the spot market at that higher price. As a result, with each company generating at or very near their contract level, net trading between companies on the spot market is likely to be very minor. Therefore, while the actual level of the spot *price* has little influence on release levels, the operation of the spot *market* (including the influence of contracts) helps to determine how much generating companies will offer to generate each day.

With the introduction of a release model that accurately models market hydro reservoir storage trajectories, the combined price and release simulation model is now able to forecast and backcast over a wide range of inflow sequences. The most important consequence of this is that it can model market behaviour in periods where no market actually existed, as long as a sequence of inflows exists. This makes the combined model very powerful, and able to be applied in many different hypothetical and historic situations, several of which are explored in the following chapter.

[21] Generating an amount greater than their contract level will decrease the spot price, assuming the company has market power, however this may still be profitable as long as the marginal revenue they receive from this generation is greater than their marginal cost of generation. Although the spot price falls, the increase in revenue may still be sufficient to generate.

7

APPLICATIONS OF THE HYDRO SIMULATION MODEL

7.1 Introduction

As explained at the conclusion of the previous chapter, the price and release models developed in this thesis may be combined into a Monte Carlo simulation model of New Zealand's aggregate storage levels and daily average spot prices, requiring as input only a starting storage level and an inflow sequence. This leads to several interesting applications of the model, some of which are detailed in this chapter. For example, in Section 7.2 two well-known hydrological events in New Zealand's electricity history are examined in more detail to identify how the market might have been "expected" to behave, given the hydrological conditions observed. In Section 7.3, comparisons are made between expected market behaviour and the storage behaviour observed under the different regimes since 1980. Finally, in Section 7.4 a "long-run" price duration curve (PDC) is estimated, showing that prices in the 1999-2003 period have been, on average, higher than would be expected in the longer term. The implications of this finding are discussed at the conclusion of the chapter.

The results presented in this chapter should be viewed with some prudence, however. Any forecasting or backcasting over a specific time period using a model whose parameters were estimated using a set of data from another time period should be examined with the knowledge that the conditions which generated the data in the two periods may be quite different. The release model attempts to estimate firms' behaviour given specific market and hydrological conditions, however in this chapter other inflow sequences are used as input to estimate how the market may have behaved given a wider range of inflows. It should be considered that demand conditions (i.e. location and profile of demand) and the supply mix are not taken into consideration in the release model, and differences in reservoir management behaviour may be due to factors that are not modelled explicitly. Further discussion of this issue is provided in later sections of this chapter.

7.2 Analysis of specific hydrological events

Releases are estimated with a dynamic regression model, which is better suited to estimating expected levels of release, rather than outlying levels. However, this attribute results in the model being particularly well suited for identifying periods in the past in which releases (and release behaviour) differed significantly from expected levels. While it does not account for every exogenous factor that might influence releases, the model could still find application for market regulation. It can identify periods in which generators, given their relative storage levels, are releasing less than expected, perhaps to withhold generation and increase prices, or more than may have been expected, perhaps to reduce their storage and increase prices at a later date.

7.2.1 The 2001 water shortage

An obvious example of a period to which the model can be applied is the winter of 2001. It is interesting to examine the release behaviour in the time leading up to and around the period of low inflows that caused prices to increase in the middle of the year.

As at 1 January 2001, the aggregate national storage level was 3899 GWh, the highest at any time during our sample period, and the highest it had been since November 1998. From the point of view purely of storage, there certainly was not any reason to believe that there would be a shortage of water later that year.

Median simulated releases[1] for the month of January 2001 (starting at a storage level of 3899 GWh) would total 2481 GWh. However, the total releases observed for that month were 2664 GWh. There had been a nett inflow of 1034 GWh in December 2000 due to high inflows, but in January 2001 there was a nett decrease of 572 GWh; this at a time when storage is usually being filled rapidly for the coming winter. While it is possible that some of the reservoirs were nearing their capacity at this stage and required emptying, the Waitaki reservoirs were still far from full. This rapid decline in storage can be seen in the first month of Figure 7.1 below.

[1] Unless otherwise stated, 1000 independent simulations were conducted to calculate the median releases and storage trajectories.

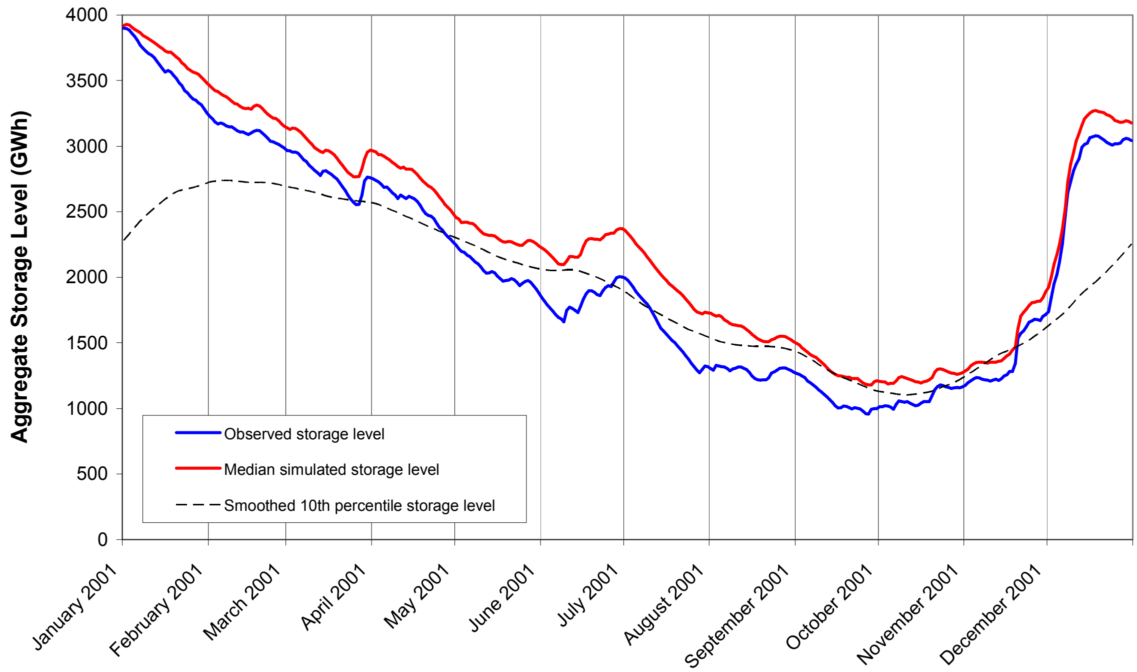

Figure 7.1: Aggregate New Zealand observed storage level and median simulated storage level, 1 January 2001 – 31 December 2001

Lower than average inflows for the first four months of 2001 (see Figure 7.2) led to the steady decline of the storage level from the relatively comfortable position at the start of the year toward the historic 10th percentile level. In May, with storage approaching dangerously low levels, releases were again unusually high – this time about 9.5% greater than simulated releases for the month. This can be seen in the fact that the two storage trajectories in Figure 7.1 are not parallel for that month.

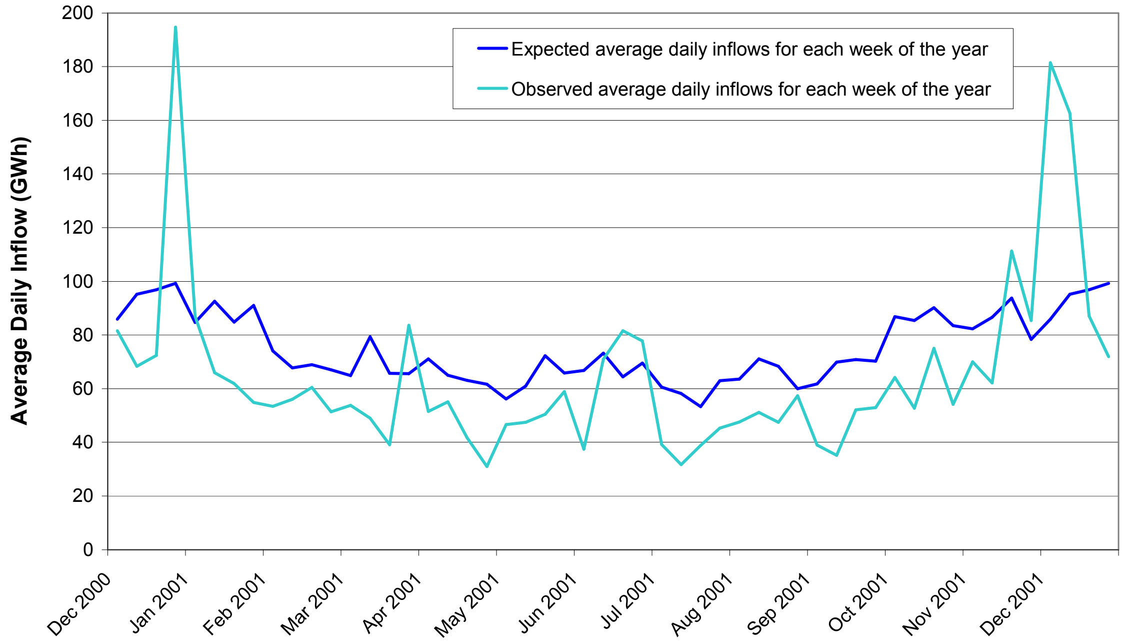

Figure 7.2: Expected aggregate daily average inflows for each week of the year, (calculated over the whole 1980-2003 sample period) and observed average daily inflows for each week, 1 January 2001 – 31 December 2001

Actual releases match modelled releases in June and July, with a more normal level of inflows in June boosting storage and causing a temporary decrease in the spot price, as can be seen in Figure 7.3 below.

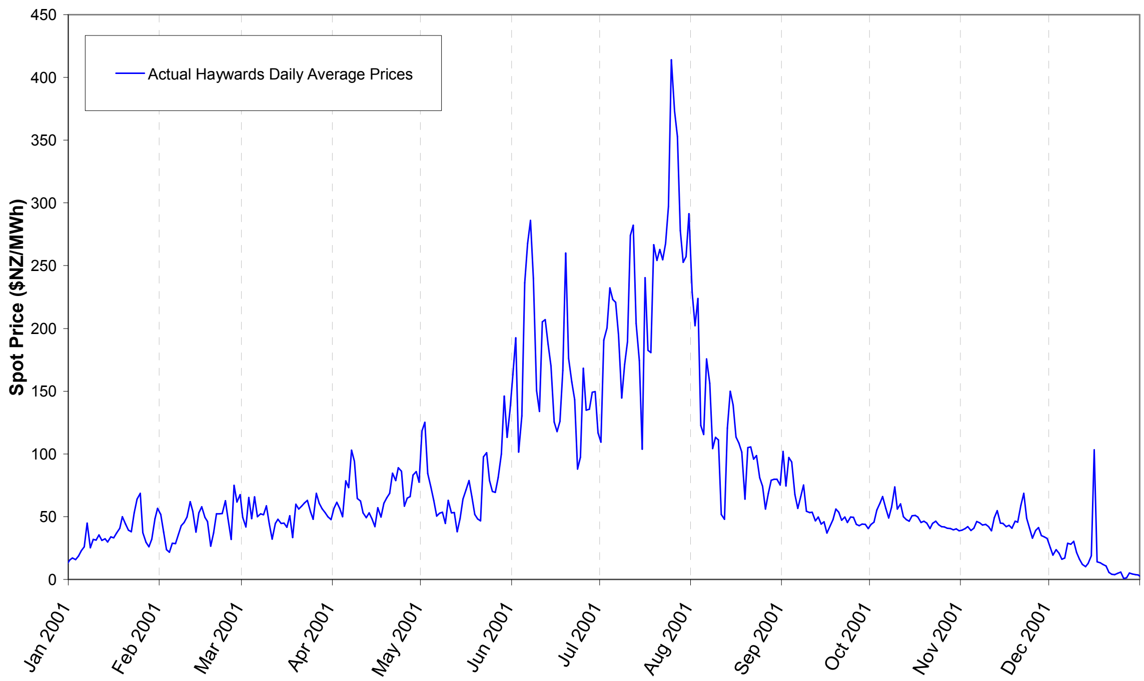

Figure 7.3: Daily average spot prices from the Haywards node, 1 January 2001 – 31 December 2002

Very low inflows in July (see Figure 7.2) raised spot prices to levels even greater than they had been in June. However, the effects of a widespread public electricity conservation campaign began to be felt in August. Despite inflows still much lower than average, the actual storage level was maintained at a constant level throughout the month, helping to decrease spot prices somewhat. By way of computation, the simulated storage level would have been much higher than actual storage at this stage, allowing modelled releases to be 10.5% greater than actual releases.

Inflows were again substantially lower than average in September and October, causing the storage level to remain at grave levels until very high inflows from mid-November onwards stabilised the situation. It is interesting to note that the sharp drop in spot prices in August apparently had very little to do with the storage level, which could not be considered "safe" until December at the earliest. Despite the insecurity of supply caused by the low storage levels, prices remained at normal levels for the rest of the year.

One factor possibly responsible for the unusual storage and price behaviour observed in 2001 was the contracting situation of electricity retailer On Energy (OE). OE did not sufficiently hedge its spot price risk leading into the winter of 2001, and suffered heavy losses on the spot market during the period of high prices. This led to OE exiting the retail market; their North Island customers were acquired by Genesis Energy Limited, and Meridian Energy Limited (MEL) acquired the South Island customers.

It has been suggested that OE failed to renew its normal contracts with MEL early in the year. If so, there were effectively two changes in industry structure within 2001, with the first occurring when OE failed to renew its normal contracting arrangements, thus effectively reducing the contract exposure of its suppliers, principally MEL, by the same amount. Thus, the observation that releases seemed higher than expected at that time of the year, would be consistent with the suggestion that potential suppliers believed they would no longer need to retain water to supply OE's customers over the winter, or at least that they no longer had any legal obligation or commercial requirement to do so.

Being vertically integrated, Genesis and MEL then became more heavily contracted when OE exited the market, and it would have then become in their interests to offer generation to the market in a way that would decrease the spot price. Whether or not this actually happened could only be determined by examining the historic offer stacks of each of the two companies from this period. What is interesting to note is that the 2001 power crisis (and hence the public savings campaign) was deemed to be "over" when the spot prices dropped, and not when storage returned to higher levels. Again, inflows were low in September and October, therefore it is unlikely that a forecast of impending high inflows was responsible for the decrease in prices.

As noted earlier, the public electricity savings campaign helped to maintain storage levels in August 2001 at a constant level in the face of much lower than average inflows. The simulated storage level in Figure 7.1 is not nearly as low at the start August as the actual level, and hence the urgency to reduce releases would not have been as great. It is an interesting exercise, however, to begin the simulation at the start of August and study the actions of the release model starting from such a position. As Figure 7.4 shows, storage (and hence implicitly releases) are matched nearly perfectly for the whole month of August, when the market was in "savings mode", and for some of September. The release model acts exactly as the market did. However the modelled releases again underestimate actual releases from the middle of September. It could be speculated that actual releases increased again at this point due to the fact that spot prices decreased, the power crisis was deemed to have ended, and demand increased again after the savings campaign finished. Alternatively, it may be suggested that at this stage the market returned to "normal mode", which it had been in before OE failed to renew its contracts with MEL. Before this occurred, MEL may have had a higher level of retail contracts, as it did again after the exit of OE. In between these two events they were under-contracted, and hence their incentives to release would have been significantly different.

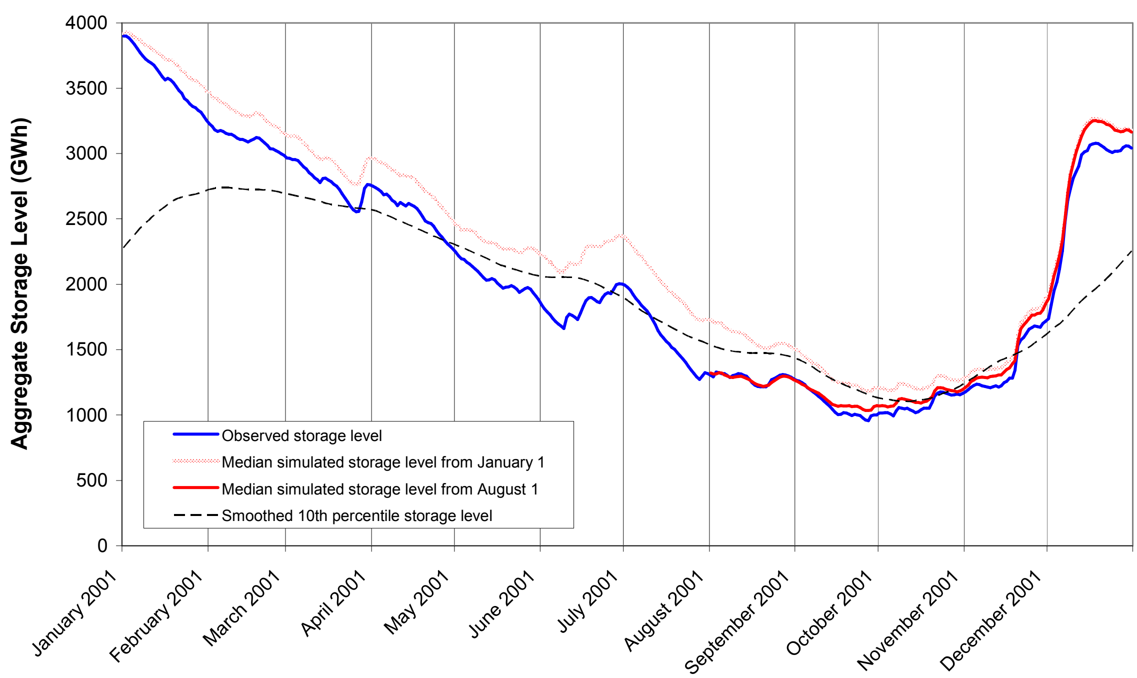

Figure 7.4: Aggregate observed storage level and aggregate median simulated storage level, 1 January 2001 – 31 December 2001, and 1 August 2001 – 31 December 2001

7.2.2 Storage levels in the pre-market era

While much has been made in recent times of the periods of low inflows in 2001 and 2003, using a wider data set helps to put those periods in perspective[2]. Figure 7.5 shows the observed and median simulated storage trajectories, starting from January 1980 and continuing until June 2003, given the actual inflow sequence that occurred.

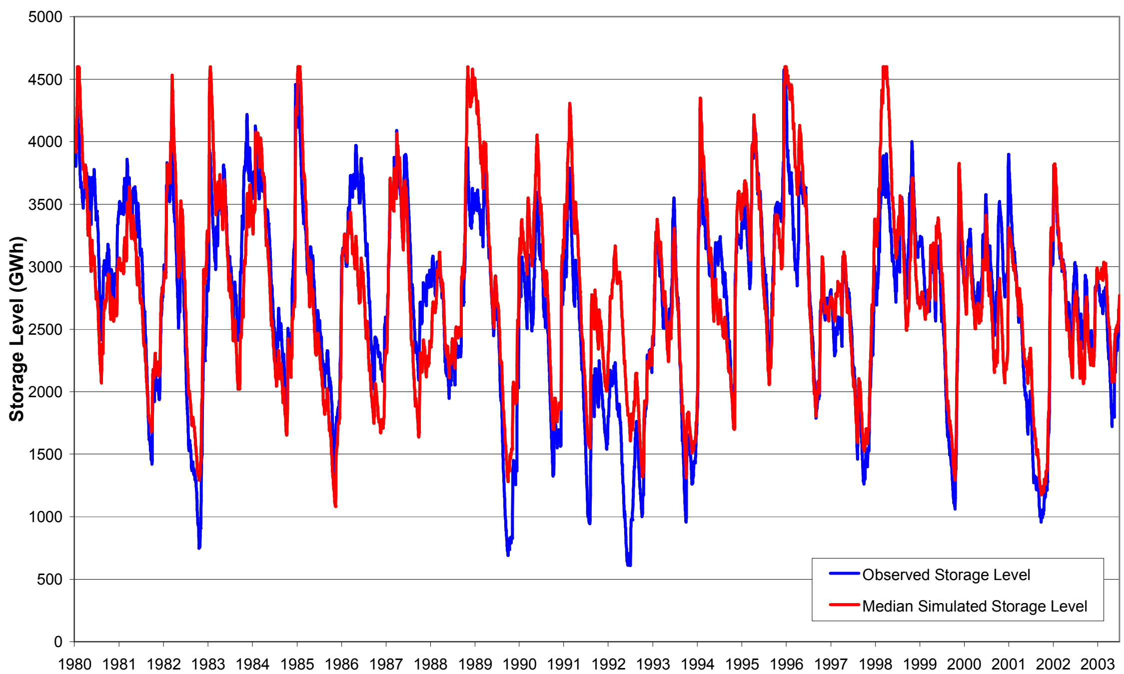

Figure 7.5: Aggregate observed storage level and aggregate median simulated storage level, 1 January 1980 – 30 June 2003

One thing that is immediately apparent is that the simulated storage level is overestimated in some of the summer peak periods, though not after 1999. A reason why this occurs may be that the 1999-2003 data, from which the parameters of the model are estimated, contains no periods of sustained high inflows (i.e. an average of over 115 GWh per day

[2] Note that some discussions in this chapter rest on the assumption that 1980-2003 provides a more representative sample of hydrologies than was observed in 1999-2003. Different conclusions may apply if the 1980-2003 period was unusual, or the climate is changing (some commentators have suggested that the next 20 years may actually be dryer).

for more than two months), which occurred each of the five or so times from 1980 that the simulated storage trajectory is too high. At times of high inflows, release behaviour is a great deal more variable and may not be well modelled by a dynamic regression model for expected releases. Also, as with any model based on exogenous data, caution should also be exercised when estimating values of a dependent variable using data that may contain values outside the range over which the parameters of the model were estimated. In this case, the release model was estimated using a dataset that contained several periods of low storage levels and extreme dry inflow sequences, but no extreme wet sequences coupled with high storage levels.

To counter this, the simulation model can be programmed to spill water when aggregate storage is greater than 4600 GWh (the highest amount recorded in between 1980 and 2003). In reality, it is practically impossible for every single reservoir in New Zealand to be full at the same time. Therefore, in practice it is possible for the flows into some of the reservoirs to be spilled even when the aggregate storage level is fairly low. This is hard to model in a one-reservoir model, hence the need for a more general rule for spill.

Another notable anomaly in the simulated data is the fact that actual storage was maintained at somewhat lower levels than the simulated storage during the time that the Electricity Corporation of New Zealand (ECNZ) provided the country's electricity generation, from 1987 to 1996. This period included 1992, which saw the most significant and memorable power crisis in New Zealand's recent electricity history. At the time, speculation abounded as to the reasons why the storage was run so low; the release simulation model is able to examine the situation from a different angle. Further investigations into the differences in storage behaviour between different regimes is provided in Section 7.3.

As mentioned at the start of this chapter, the combined price and release simulation model takes no account of the nature of demand (location and profile) or of the supply mix. The location and daily profiles of demand for electricity in New Zealand have changed since the early 1980s, with the majority of the population now living in the northern third of the country and new energy-intensive forms of industry such as dairy farming changing demand conditions within the country. Furthermore, the ratio of hydro to non-hydro generating capacity has decreased over the same period as more thermal plant have come on line, and the amount of hydro generating capacity (with significant storage) has remained largely constant since the commissioning of the Waitaki system. These conditions will obviously change the release behaviour of a firm or Government that is optimizing its strategy; a release policy that is optimal given one set of conditions will not necessarily be optimal when those conditions are changed. This should be considered when examining the results in the remainder of this chapter.

7.2.3 The 1992 storage crisis

It is clear from Figure 7.5 that the observed storage level in the winter of 1992 was lower than at any other point since 1980. What is also clear, however, is that the storage was also very low in the winter of 1991, and did not recover sufficiently during the summer in between. This is more clearly displayed by the blue line in Figure 7.6 below, representing the observed storage level. The red line shows that, starting from the same storage level on 1 February 1991, the simulated storage trajectory would also have reached significantly low levels in both of the two years. Interestingly, while the storage level was relatively high immediately after the winter of 1991, even the simulated storage trajectory reaches relatively low levels at the end of that year.

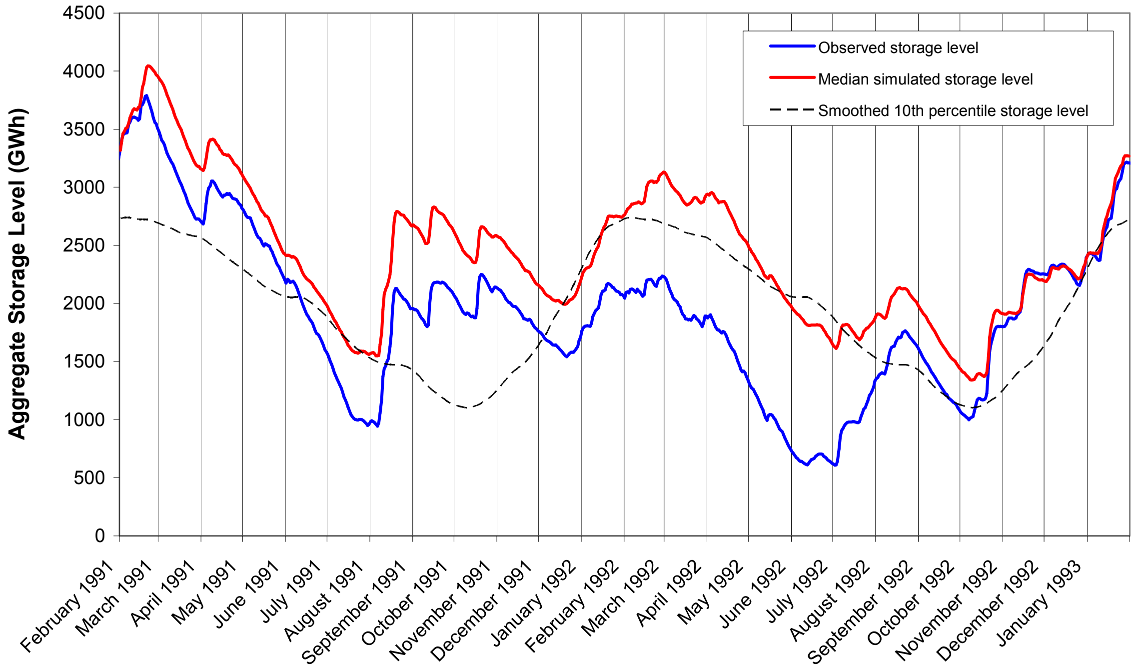

Figure 7.6: Aggregate observed storage level and aggregate median simulated storage level, 1 February 1991 – 31 January 1993

The sequence of inflows responsible for these storage levels is shown in Figure 7.7. The three months over summer of 1991-2 were very dry, and significantly lowered the reservoirs at a time when they required refilling.

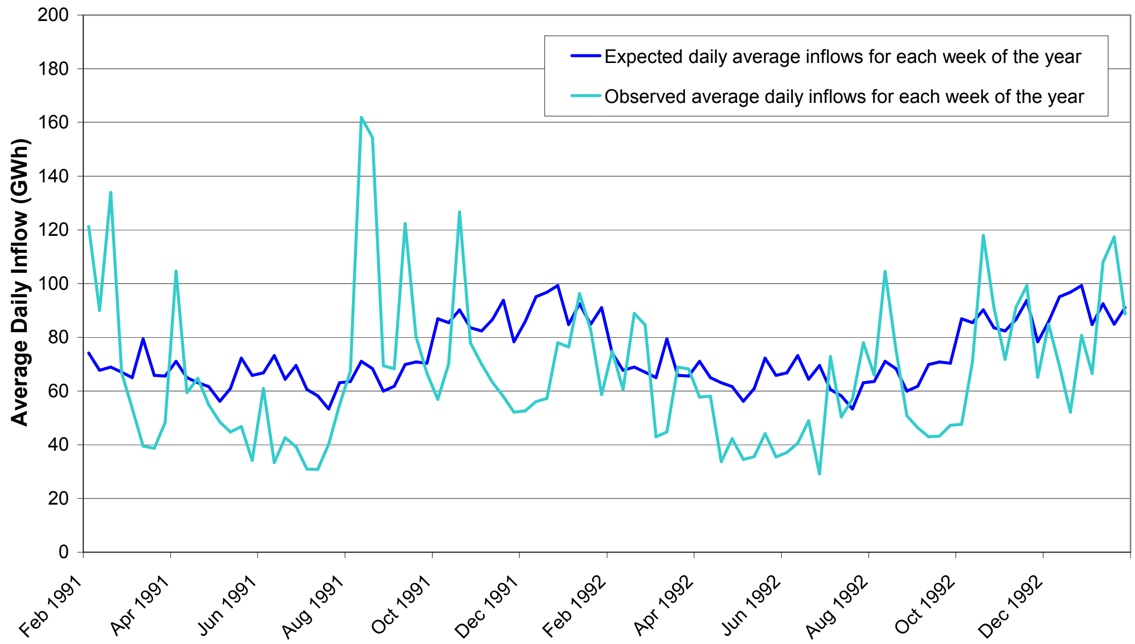

Figure 7.7: Expected daily average inflows for each week of the year, (calculated over the whole 1980-2003 sample period) and observed average daily inflows for each week, 1 February 1991 – 31 January 1993

It is interesting to note from Figure 7.6 how the market may have responded given the same sequence of inflows. The median simulated storage runs to a relatively lower level in 1992 (around –250 GWh) than it does in 2001. The "market" (as represented by the release model) would not have run storage down to anything like the low levels that actually occurred in 1992, particularly around June. The green line in Figure 7.8 below illustrates this point further; given the very low starting storage level in January 1992, initial release levels would have been very low to enable the reservoirs to fill to a reasonable level. Once the simulated storage level recovers to a suitably high level, the simulated storage trajectory then runs parallel to the actual trajectory until approximately June or July.

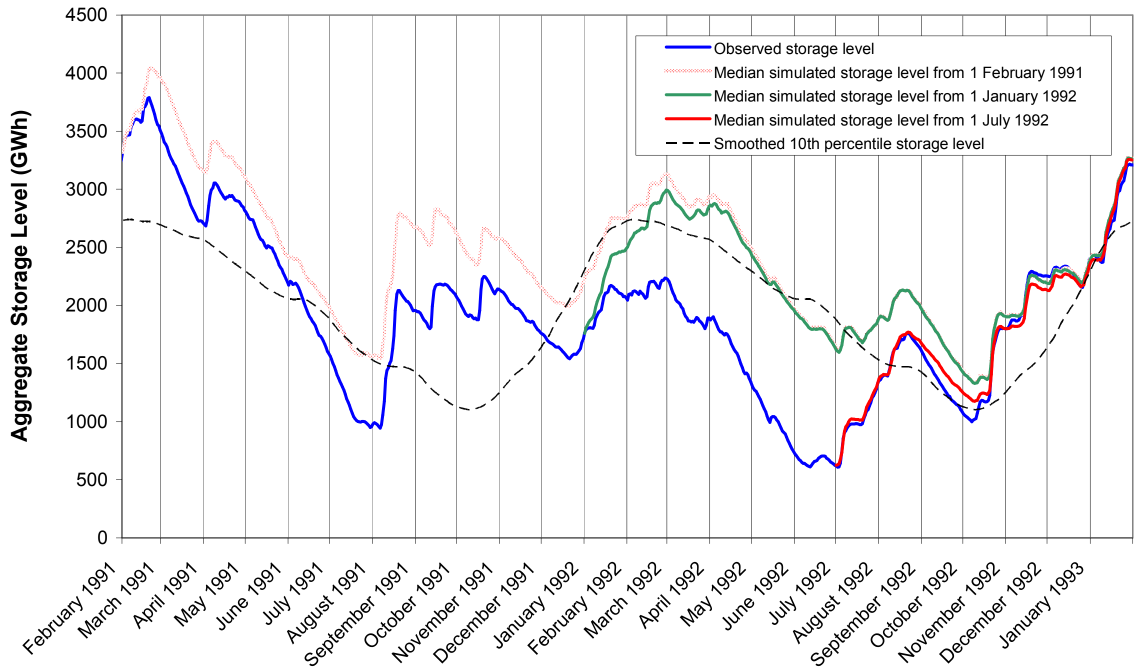

Figure 7.8: Aggregate observed storage level and aggregate median simulated storage level, 1 February 1991 – 31 January 1993, 1 January 1992 – 31 January 1993 and 1 July 1992 – 31 January 1993

With the national storage level as low as it was in July 1992, and in danger even of running out completely, a vigorous public electricity savings campaign was launched, similar to that which occurred in 2001. Whereas the average daily release in June and July over the whole length of the 1980-2003 sample period was 71 GWh, in June and July 1992 observed releases dropped to an average of 42 GWh per day. Coupled with above-average inflows in July and August (see Figure 7.7), the result of the power savings campaign was that storage levels rose quickly and were much safer by November, despite a very dry September. What is interesting to note is that, as in the 2001 case, the simulated releases match very accurately the actual releases in the face of a power savings campaign (see the thick red line in Figure 7.8). We conclude from this that the fitted market model is particularly adept at rescuing itself from situations where storage levels are (or are becoming) dangerously low, even if it would not have released so much as to have fallen into those situations to begin with. This suggests that the market would have remedied the storage crisis with MWV-controlled releases and high spot prices, whereas the Government-controlled generator of the day could only achieve this with the aid of a major public savings campaign.

7.2.4 The relative severity of various low inflow sequences

After examining the effects of two particular sequences of low inflows, the question remains: just how extreme were these sequences, and which would have led to the most extreme price outcomes in a market era? Obviously, the lower the relative storage level (RSL), the more serious the water shortage and the higher the market prices would be. Therefore, we need to examine the simulated relative storage levels and market prices from 1980 – 2003 to answer the question. The results are shown in Table 7.1.

Period	**Median Simulated Relative Storage Level**
Dec 1993 – Jan 1994	-460 GWh
Nov 1985	-361 GWh
Jan 1987	-250 GWh
June 1992	-248 GWh
April – May 2003	-239 GWh

Table 7.1: The five most extreme periods of simulated relative storage shortage observed under "market" conditions from January 1980 – June 2003

While the lowest RSL actually observed over the 1980-2003 period occurred in 1992 (-1443 GWh), the most serious shortage in terms of simulated market outcomes would have been at the very start of 1994. Under the assumption that the exponential relationship between spot price and the RSL holds for all levels of the RSL, then the spot price at the end of 1993 would have been more than twice as high as it was in 2003. The simulated and actual storage trajectories from the 1993-4 period are shown in Figure 7.9 below.

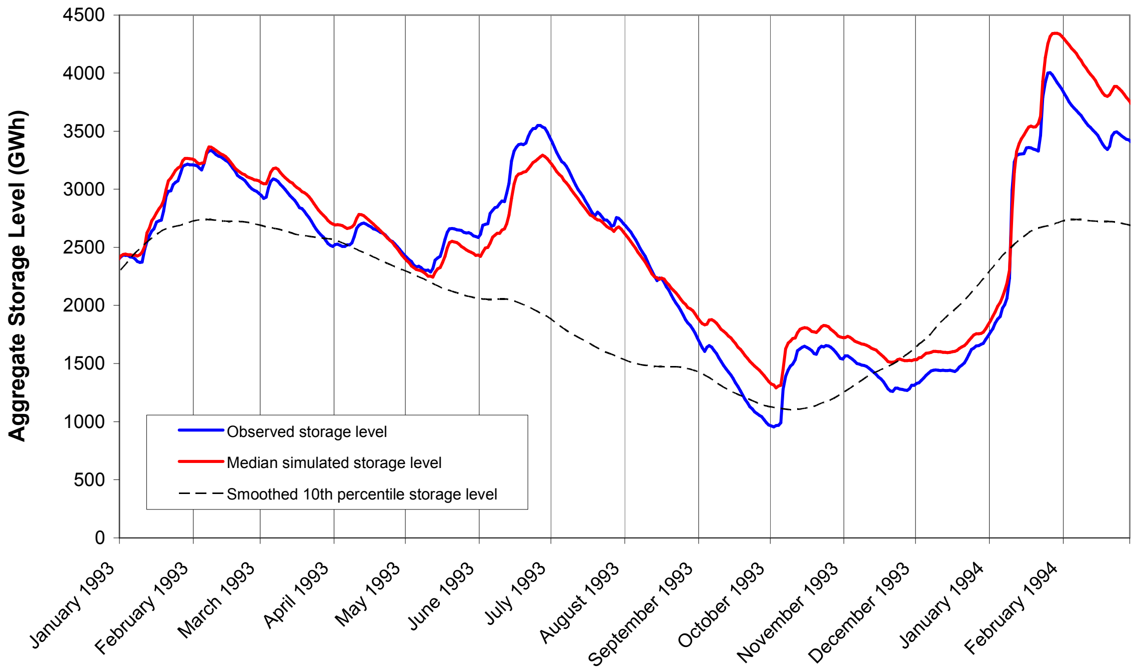

Figure 7.9: Aggregate observed storage level and aggregate median simulated storage level, 1 February 1991 – 31 January 1993

The storage trajectories show that this situation was not necessarily a direct result of the dry year in 1992, as the RSL was not particularly low in July 1993. Very low inflows from July through to December resulted in the end-of-year storage level being dangerously low. Were it not for extremely high inflows from the second week of January 1994, the situation could have been considerably more serious the following year. The extreme inflow sequence that led to this situation is shown in Figure 7.10.

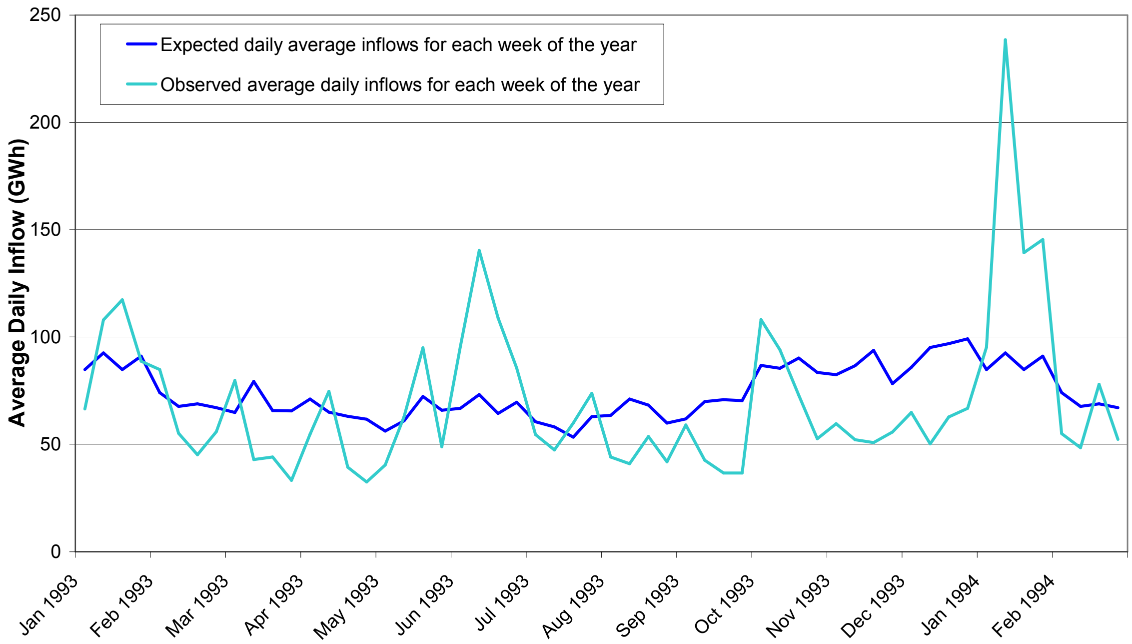

Figure 7.10: Expected daily average inflows for each week of the year, (calculated over the whole 1980-2003 sample period) and observed average daily inflows for each week, 1 February 1991 – 31 January 1993

As such a period of low inflows in summer did not occur in our sample period, it is impossible to know exactly how prices would respond[3], however we could presume that market prices would have been very high at the start of 1993. Had the high inflows of January 1994 been forecast, however, then the actual market price level may not have been high as the price model, which does not take into account any forecasts, would suggest.

7.2.5 Estimated behaviour with a hypothetical extreme inflow sequence

An interesting exercise would be to see just how low storage levels would be run in the market era if an inflow sequence was constructed from the lowest observed inflow

---

[3] Recently, however, the RSL was low in the 2005-6 summer period and spot prices did average above $200/MWh for some weeks. Recalibrating the model with data including this period would enable comparisons between the 1993-4 summer inflows and those of 2005-6.

months in the 1980-2003 period (i.e. the lowest January followed by the lowest February followed by the lowest March etc.). If this inflow sequence (shown below in Figure 7.11) were to repeat itself for more than one year, would storage run dry?

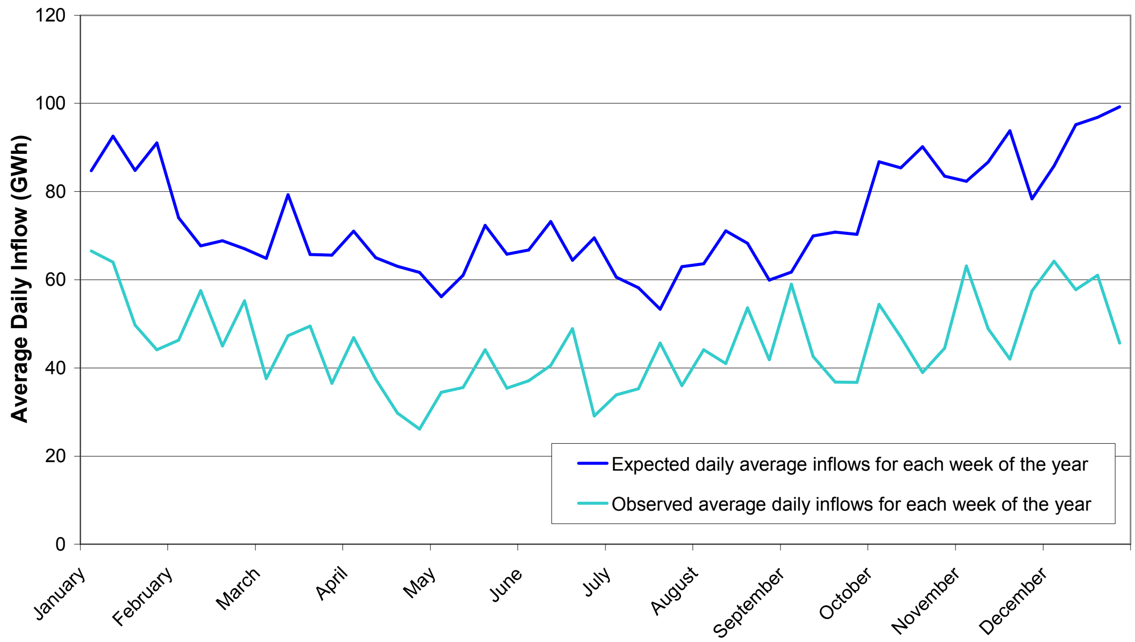

Figure 7.11: Expected daily average inflows for each week of the year, (calculated over the whole 1980-2003 sample period) and lowest observed complete months of inflows over the 1980-2003 sample period

The median simulated storage trajectory in Figure 7.12 answers this question. Starting from a beginning-of-year storage level of approximately 2500 GWh, storage would not actually be run dry, however releases would have to be reduced significantly in order for total release in a year to equal total inflows. This would result in a significant rise in the Loss of Load Probability[4]. The RSL would hit a minimum of -850 GWh at the end of December, resulting in extremely high spot prices and very low releases. In fact, releases would be so low that there would be a high chance of black-outs; average daily releases for June – August are usually around 70 GWh, but in this case they would be only 48

[4] The Loss of Load Probability is the probability that system demand will exceed generating capacity, resulting in power blackouts.

GWh – about 70% of usual hydro generation. As hydro now accounts for approximately 65% of all New Zealand's usual electricity generation, this would mean that an extra 20% of total load would need to be met by non-hydro generation.

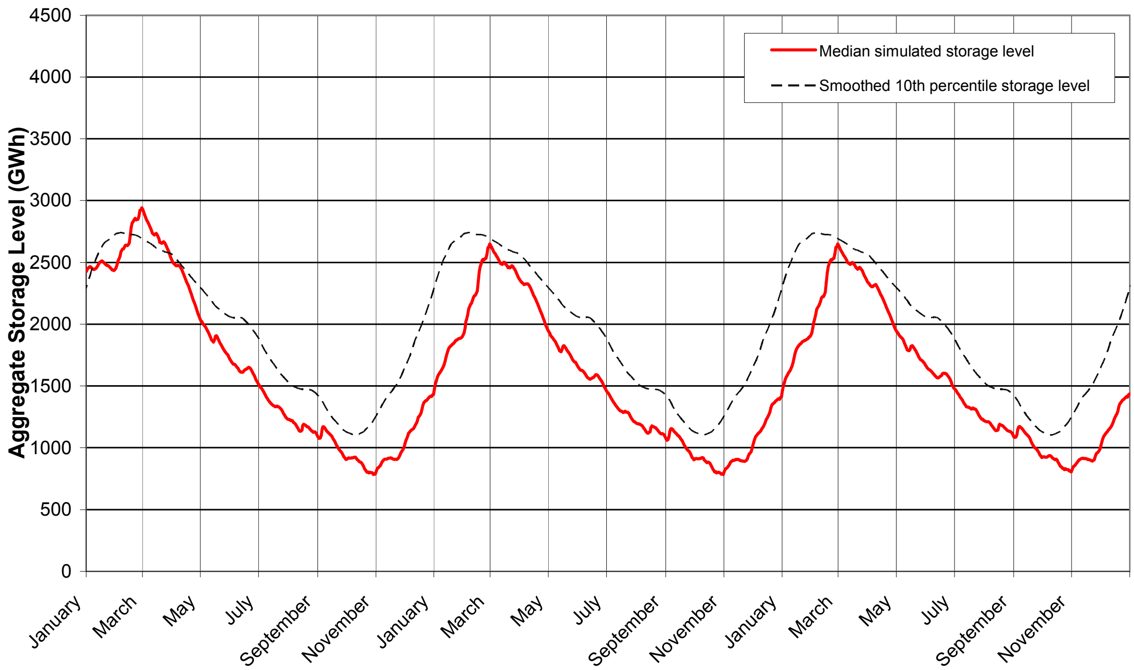

Figure 7.12: Aggregate observed storage level and aggregate median simulated storage level, 1 February 1991 – 31 January 1993

7.3 Comparison between different storage regimes

Broadly speaking, from 1980-2003 New Zealand's hydro reservoirs were operated by three separate regimes. The Ministry of Energy (MoE) was responsible for the generation and supply of power until 1987, when the Electricity Corporation of New Zealand (ECNZ) was established[5]. ECNZ operated as the sole generator and supplier of electricity in the country until the New Zealand Electricity Market (NZEM) was formed late in 1996. But the current market regime can best be characterised as commencing with the break-up of ECNZ in 1999.

[5] The New Zealand Electricity Department (NZED) operated as a separate Government department for many years before the MoE was formed. This analysis could be extended to cover those years.

Thus the combined price and release simulation model, calibrated using market data from 1999-2003, can be used to estimate how the market may have behaved given the same inflow sequences as the MoE and ECNZ experienced. The results in this section were generated by resetting the storage level at 1 January each year to the level that actually occurred, rather than generating continuous sequences from 1980-2003. This ensures independence between simulated years, as the simulated results in one year in no way affect the results in any other year, and allows for direct comparison between the actual storage trajectories observed each year and those simulated by the model. Two hundred storage trajectories were simulated for each year[6], and the results were aggregated into three periods: MoE (1980-1986), ECNZ (1987-1996) and NZEM (1997-2003). The results analysed in this section include, for each year, the differences between actual and simulated average *storage* levels, between actual and simulated *minimum* storage levels, between actual and simulated *end-of-year* (EOY) storage levels, and in the timing of actual and simulated minimum storage levels.

7.3.1 Difference in annual average storage levels

It is worthwhile comparing the difference between actual annual storage levels and simulated annual storage levels, as it can confirm whether or not, on average, storage was run down any harder in any particular regime, as suggested in Figure 7.5. It is generally believed that storage levels were low in the ECNZ years, but this may have been due entirely to the hydrology in those years, rather than a particular change in behaviour.

Figure 7.13 below shows the frequency of the differences between annual average simulated and actual levels for each regime. In the market years since 1997, simulated

---

[6] These are not 200 different hydrology sequences, but 200 different realisations of the stochastic process describing the way in which market outcomes can be expected to deviate from the "deterministic" (but still hydrology-dependent) process fitted by the model, as a result of factors other than hydrology. The hydrological sequence is the same for all simulations, being the sequence actually observed for the year in question.

average storage levels are, overall, no different from actual annual storage levels. Had there been a discrepancy between these two results it would not have been surprising, given that the model was estimated from 1999-2003 data, not 1997-2003. This graph shows that, on average, storage levels were kept at lower levels during the ECNZ years than the market would have kept them, but the market would have run storage levels lower than the MoE did. Thus, in this respect, the behaviour of the market regime lies between these two other regimes.[7]

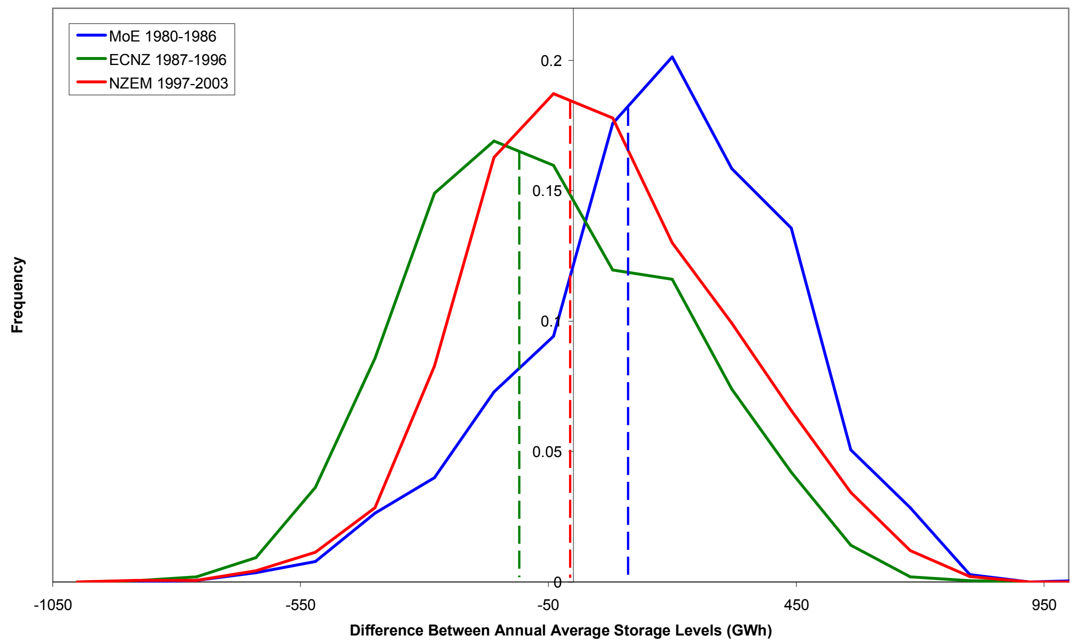

Figure 7.13: Distributions of the differences between actual annual average storage levels and simulated annual average storage levels, 1980-2003. Averages of each distribution are given by dotted lines

[7] A statistical hypothesis test performed on these simulation results (see Appendix G) suggests that there is enough evidence at the 5% level of significance to conclude that while the mean difference in the NZEM era is not significantly different from zero, the mean differences in the other eras are. There is also enough evidence to conclude that the mean difference in the ECNZ era is less than that of the NZEM era, and the mean difference in the MoE era is greater than that of the NZEM era. It should be recognised, though, that this really only says that a model fitted to the NZEM years produces simulation results for other eras which are statistically different from the actual average levels observed in those years, given the uncertainty introduced by the random stochastic elements of that model, assuming fixed hydrology, as in these simulations. It does not prove that there is a statistically significant difference between the real levels in both eras.

7.3.2 Difference in annual minimum storage levels

Comparing the minimum actual and simulated storage levels each year offers a second assessment of whether storage is run down any lower in the market era compared with in other regimes. Figure 7.14 shows the distribution of differences between actual annual minimum storage levels and simulated annual minimum storage levels. The graph shows that, on average, the actual minimum storage levels are greater than simulated storage levels in both the market and MoE era.

This suggests that the model, which was calibrated on 1999-2003 data, may not be perfectly calibrated with respect to this aspect of performance for the 1997-2003 period. But the discrepancy is not large. More importantly, it suggests that, starting from the same storage position, the market regime draws down storage to a greater degree than that which occurred during the MoE years. Thus there appears to be no evidence of any systematic tendency by the market to withhold water, or more exactly hydro power, to raise prices or for any other reason, over the winter season.

The actual minimum storage levels are somewhat lower in the ECNZ era than in the other two. In fact they are very close to the levels simulated by the model as calibrated for the 1999-2003 period. This is consistent with the belief that ECNZ tended to run reservoirs down somewhat more than had previously been the case. But this result is not really conclusive, particularly because both wet years and dry years were included in the simulation, and storage behaviour is less defined in wet years than dry years. The results may have been different if the minimum storage levels in dry years had been compared, but that would have meant drawing inferences from a very small sample, with essentially arbitrary judgments being made as to which years were "dry", and which were not. [8]

---

[8] In this case there is enough evidence at the 5% level of significance to conclude that the mean differences in each of the eras are significantly different from zero. Again, there is also enough evidence to conclude that the mean difference in the ECNZ era is less than that of the NZEM era, and the mean difference in the MoE era is greater than that of the NZEM era.

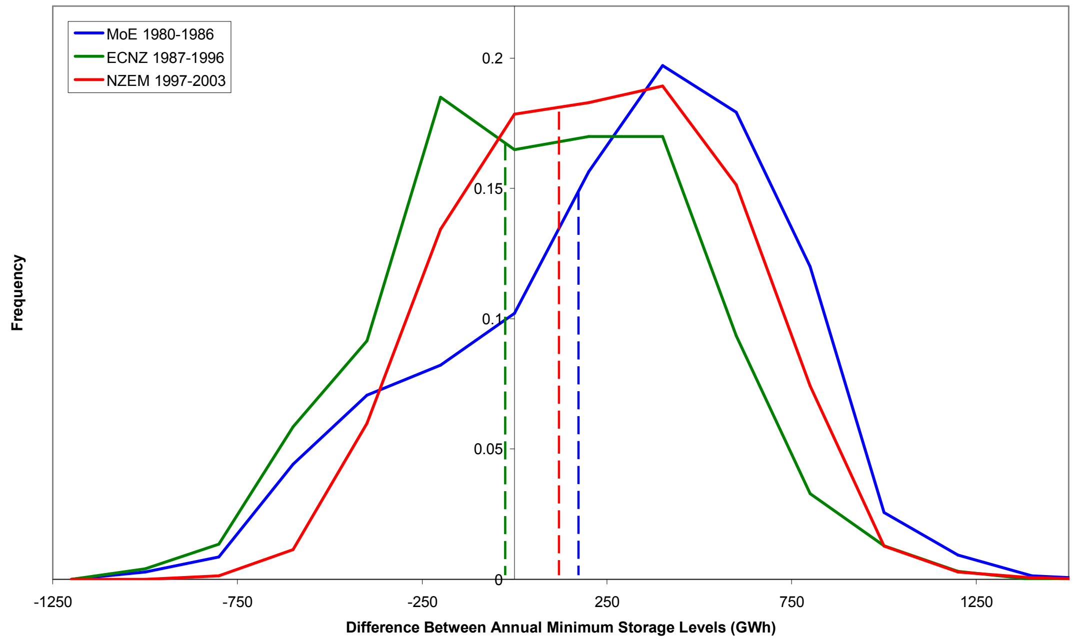

Figure 7.14: Distributions of the differences between actual annual minimum storage levels and simulated annual minimum storage levels, 1980-2003. Averages of each distribution are given by dotted lines

7.3.3 Difference in EOY storage levels

Given that the simulated storage trajectories for each year start at the same level as the actual trajectories, it is pertinent to note where each trajectory finishes in comparison to the EOY storage level that was actually observed. Leaving a low amount of storage at the end of a year negatively impacts the following year's generating capability, which appears to have occurred more than once in the ECNZ years.

The distribution of differences in EOY storage levels is shown in Figure 7.15. This graph suggests that, on average, EOY storage was lower in the ECNZ years than the market would have left it, but higher in the MoE regime. Thus it reinforces the conclusions of the previous analysis, and suggests that the market is not only drawing reservoirs down at a similar rate to the MoE, but also re-filling them at the same rate. Thus there again appears to be no evidence of any systematic tendency by the market to withhold water, or more exactly hydro power, to raise prices or for any other reason, over the annual storage cycle.

Again, the ECNZ result is a little lower[9], which might be taken to indicate some tendency to reduce storage over an annual cycle. But such a tendency could not actually be sustainable. It should be recognised that the simulations being performed here all start from the actual 1 January storage level for the year concerned. Thus, in the ECNZ years, the simulations all start from the storage level attained by ECNZ. Thus the observed tendency to reduce storage does not indicate a tendency to lower storage over each successive annual cycle, but an essentially one-off change to a lower storage regime.

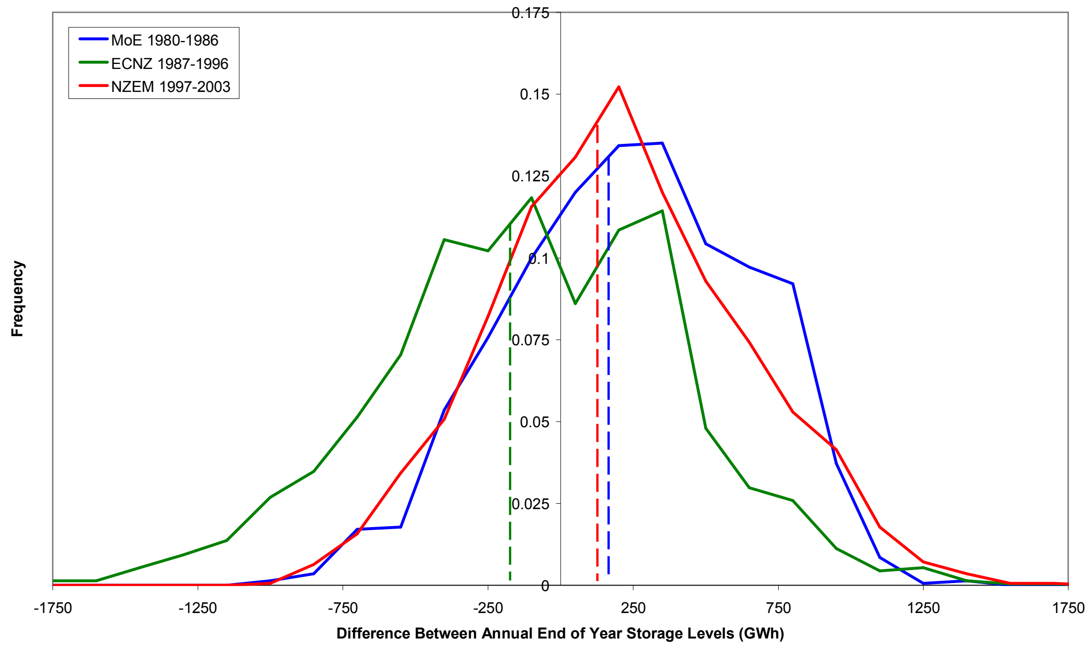

Figure 7.15: Distributions of the differences between actual EOY storage levels and simulated EOY storage levels, 1980-2003. Averages of each distribution are given by dotted lines

[9] As with the two previous sections, the results of statistical hypothesis tests on these observations reveal that, at the 5% level of significance, there is enough evidence to conclude that the mean differences in each of the eras is significantly different from zero. There is also enough evidence to conclude that the mean difference in the ECNZ era is less than that of the NZEM era, and the mean difference in the MoE era is greater than that of the NZEM era.

7.3.4 Difference in the timing of minimum storage levels

Given that each simulated storage trajectory starts at the same level as was actually observed that year, and New Zealand generally has a regular seasonal pattern to its hydro storage, comparing the times at which the minimum simulated and actual storage levels occurred offers some idea as to whether storage is being drawn down at a lesser or greater rate in the NZEM era. That is, it could indicate whether there is any systematic change to the pattern of draw-downs, rather than just the total draw-down amount.

Table 7.2 shows the dates at which the minimum storage levels actually occurred, and the median dates at which the simulated minimum storage levels occurred. These results are conclusive; apart from three years in the ECNZ era, 1988, 1992 and 1994, the difference in timing between actual and simulated minimum levels in each year is no more than fifteen days. This suggests that while storage may have been drawn down a little more in the ECNZ years, and perhaps a little less in the MoE years, when compared with the NZEM, there has been no discernible change in the timing of this drawing down period[10]. Thus any differences in the amount drawn down can reasonably be interpreted as also indicating differences in the rate of draw-down.

[10] The differences between dates of maximum storage were also considered. But, while New Zealand storage is generally drawn down to a minimum point only once, within a fairly tight window in the middle of the year, the range of dates over which storage could be at its maximum level is greater. Comparing the timing of maximum storage levels seems unlikely to lead to any further conclusions than can already be gleaned from the information regarding minimum storage levels, minimum storage dates, and EOY storage levels.

Year	Actual	Median simulated	Days difference
1980	23 August	13 August	11
1981	26 September	26 September	0
1982	23 October	24 October	-1
1983	15 September	20 September	-4
1984	5 October	11 October	-5
1985	16 November	17 November	0
1986	18 November	24 November	-6
1987	4 October	1 October	3
1988	1 July	9 June	22
1989	6 October	4 October	3
1990	23 October	12 October	12
1991	10 August	7 August	4
1992	5 September	10 October	-34
1993	10 November	2 December	-21
1994	1 May	2 January	120
1995	18 August	18 August	1
1996	7 September	5 September	3
1997	14 September	29 September	-15
1998	22 September	17 September	6
1999	20 October	21 October	-1
2000	2 October	1 October	1
2001	10 October	5 October	5
2002	15 August	13 August	2

Table 7.2: Differences between the dates of minimum storage actually observed, and the median dates of minimum storage observed

7.4 Calculation of the Long Run Market Price Duration Curve

The original price model enabled prices to be calculated using storage levels since the market started in April 1999. Hypothetical market prices could be calculated using the observed relative storage levels pre-1999; however, that would not produce an internally consistent result because the historical storage levels reflect historical release policies, rather than those of the market. As the release model is able to produce "market" storage trajectories given actual sequences of inflows, prices can then be calculated from the relative storage levels corresponding to these trajectories.

Calculating market prices over a longer period of time will enable the high prices of 2001 and 2003 in particular to be put in the context of a longer timeframe. This will show whether the observed market prices were higher than they should have been given the actual inflow sequences that occurred, or whether the high prices simply reflected an extraordinarily low level of inflows. Running Monte Carlo simulations over both the periods April 1999 – June 2003 and January 1980 – April 2003 gives the simulated price duration curves shown below in Figure 7.16.

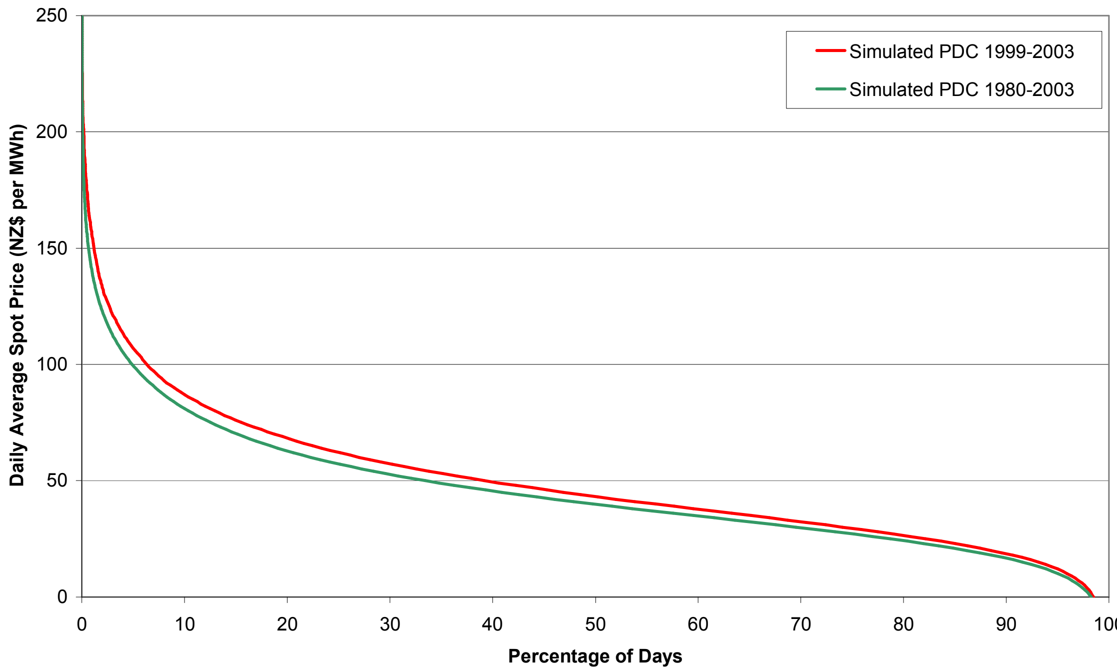

Figure 7.16: Simulated PDC from April 1999 – June 2003 (red line) and simulated PDC from January 1980 – June 2003 (green line)

The simulated PDC over 1980-2003 could be thought of as estimating the underlying PDC for the market years, from which the actual observations were sampled. This chart shows that, on average, market prices were higher from 1999-2003 than could be expected on average[11]. This is because in two of those five years there were

[11] But note again that, while a historical comparison is useful for assessing market performance, any conclusions about long run projections rest on the assumption that 1980-2003 provides a representative sample of hydrologies.

extraordinarily dry inflow sequences, which led to lower than average storage levels. What is not exactly clear from Figure 7.16 is that, at times, there are likely to be higher prices than have been observed so far. This is reinforced by the results in Table 7.1. They show that prices would have been far higher in several periods between 1980-1995, had the market been operating then, than they were in both 2001 and 2003.

The estimated PDC is particularly useful, in that it enables the accurate estimation of the commercial viability of capacity investment. If investors used the short run PDC observed from 1999-2003 instead of the long run PDC, they are likely to overestimate the earning ability of new plant[12]. Some commentators have asserted that investment has been insufficient, a situation which could both contribute to, and be attributed to, market power in various forms. But investment should not be occurring in response to particular dry year experiences, and would tend to be excessive if it was. Conversely, if investment is occurring at the optimal economic rate, driven by this long run average PDC, there is a distinct danger that it may incorrectly be deemed insufficient, if assessed against the actual observed PDC.

7.5 Conclusions

This concludes the analysis of hydro-based market behaviour in the NZEM. The applications in this chapter have shown that the market-calibrated approach provides insights into both observed and hypothetical situations, without the need for the assumptions required by the traditional bottom-up models.

[12] Alternatively, if they are using the observed PDC to value financial contracts, it may be that the likelihood of high prices is overestimated, and these contracts are priced higher than they should be.

8

MODELLING PRICES WITH A COURNOT MODEL

8.1 Introduction

While the majority of the price-modelling research presented in this thesis has involved incorporating physical system information into a top-down model for electricity prices, as mentioned in Chapter 1 there is another type of model that calculates prices from the "bottom up". These models require information such as the marginal costs, capacities and contract levels of generators, transmission constraints, and load, and make certain assumptions regarding the behaviour of the participants in the market. Using all this information, they form an aggregate market supply curve, and calculate the market price based on where the load (or demand curve) intersects this supply curve.

It is possible that both bottom-up and top-down models can be combined into a single type of model for modelling and forecasting spot prices. A bottom-up model will form an accurate estimate of what the underlying price level on any given day *should* be, provided that all factors influencing the price are held constant. In reality there are many factors in an electricity market that are not constant, but, in fact, vary unpredictably. Therefore, while the underlying level of prices can be simulated, the actual price may be very different from the price predicted. It is this unpredictable discrepancy between the two prices that can be modelled with a stochastic process, accounting for the random variation in prices caused by, for example, plant or transmission outages, and human error. Unlike the price models in the earlier part of this thesis, in this chapter we explore the use of a bottom-up model to estimate the deterministic (or predictable) level of prices.

One such bottom-up model is the Cournot market model, designed as a tool for modelling the price and output in oligopoly markets. In the basic Cournot model, market players choose a single level of output, and the market price is set by the level of demand at the aggregate market output level. The players choose their output based on a "best-response" (BR) function of the market demand curve and the other players' combined outputs; i.e. given a particular residual demand curve, they choose the output level that maximises their total profit. The other players in the market do likewise, and the equilibrium of the model is the point at which no single player can increase their profit unilaterally. This point is referred to as a Nash Equilibrium. There are many extensions of the Cournot model, such as including an order in which players are able to make their decisions and accounting for a price-taking fringe supply of small firms, but it is this basic form upon which we will be basing our analysis.

Cournot models have found application in a wide range of industries, and are one of the most widely-used tools in industrial organization regulation and policy-making. As such, they have found frequent application in modelling deregulated electricity markets, even before worldwide deregulation began in the early 1990s. Due to the nature of the physical characteristics of electricity, its generation and its distribution, wholesale electricity markets are often characterized by a small number of large firms and are, in general, perfect examples of oligopolies.

For many years, these models have been used to simulate both the operation of hypothetical electricity markets and the effects of hypothetical changes to existing markets. While the models simulate individual firm and aggregate market generation levels, their primary application in electricity markets, as in competition and policy analysis in general, has been to predict the level of electricity prices, which are of concern to consumers, producers and regulators. However, very little work has been undertaken to discover how successful these models were at estimating future price levels *after* either the markets have begun operation, or the changes have been implemented. The research presented in this chapter, which calibrates a Cournot model so that its simulated prices match with observed market data, attempts to increase the credibility of such models for future applications.

8.1.1 Pre-processing and post-processing

A Cournot-type analytical model that takes into account firm structure, marginal costs, generation capacities, and market power could form a reasonable forecast of deterministic price levels, provided it is calibrated appropriately. Recent work by Bushnell (2003) involved using a Cournot model to illustrate that certain generators were exercising market power in California in 2000, and his model produced monthly average prices very similar to those observed. But alone, an analytical model could never accurately estimate the higher-frequency price volatility observed in real markets. We suggest a two-step method that can be employed to ensure a better match between the simulated results and the historic results: pre-processing and post-processing.

Pre-processing involves identifying which input parameters the model's results are most sensitive to, and tuning these parameters so that they (and the model's results) better match the real-world situation. The major aim of pre-processing is to minimise the overall discrepancy between simulated prices and the market prices actually observed, to get a good fit of the *underlying* price paths. If instead the aim was to match the overall distributions of simulated and actual prices, a Monte Carlo simulation technique could be used to match market volatility as well, by randomly varying parameters such as the elasticity of demand or generating plant outage rate. Figure 8.1 illustrates the concept of the results of pre-processing. The red line shows the spot price series to be fitted, and the shaded blue region represents the component of the price that is accounted for by the deterministic model. Note that the model will be able to account for some, but not all, of the volatility in the price series. For this reason, and as spot price volatility is generally not symmetric[1], the model will produce a good estimate of the *median* price level, but will likely underestimate the *mean* price level.

Post-processing involves adjusting the actual results themselves, in order to take into account factors which the analytical model itself cannot. A Cournot model is a deterministic process, and as such it *can* account for events such as sudden increases in load, decreases in supply, and constraints in transmission that lead to price spikes, provided they can be predicted and calculated from the input variables. However, from a forecasting point of view it is almost impossible to predict exactly when a particular plant or transmission line may experience an outage; hence there will be some discrepancy between forecasted and realised prices. The discrepancy between the two series of prices in Figure 8.1 is shown in Figure 8.2. The green area represents the residual variation in prices that was not accounted for by the tuned deterministic model. As alluded to above and described in more detail in the following sections, this residual variation will not be symmetric, and will more often be positive than negative.

What these deterministic models *cannot* account for at all is volatility due to essentially subjective human factors, such as experimental strategies or human error, which may last for hours or days, if not weeks. While the design of offering strategies is becoming increasingly optimised and automated, it does still include a significant human element, therefore an extra process may be required to capture and model the extra volatility in the market outcomes that results. The post-processing method employed in this study adds a stochastic price process as an adjustment to the prices which the Cournot model

[1] While it is well documented that distributions of spot prices are generally not symmetric, as discussed in Chapter 2 the volatility of spot prices is also asymmetric. Positive price spikes are far more common than negative spikes.

calculates, thus combining the results of a bottom-up (analytical) model with a top-down (time series) model.

Overlaying a multi-period stochastic component (including a GARCH process) onto the prices simulated by a deterministic, static Cournot model raises obvious questions regarding the consistency of the modelling approach with respect to with the actual operation of an electricity market. As discussed in Chapter 2, static equilibrium models take no account of inter-temporal resource allocation or uncertainty regarding future outcomes. They also take no account of the specific constraints that generating plant face, such as ramping constraints and minimum up and down times. In reality, firms make decisions looking far further into the future than the next market clearance, in order to take account of the uncertainty and inter-temporal constraints. It would therefore be desirable to incorporate some element of randomness into any bottom-up tool used to model firms' behaviour, rather than relying on a stochastic process to account for the inter-temporal volatility. However, because of the use of a static model in the research presented in this chapter, rather than a dynamic equilibrium model, uncertainty has been excluded from the bottom-up modelling of prices. This motivates the use of post-processing, and the incorporation of a stochastic multi-period price process.

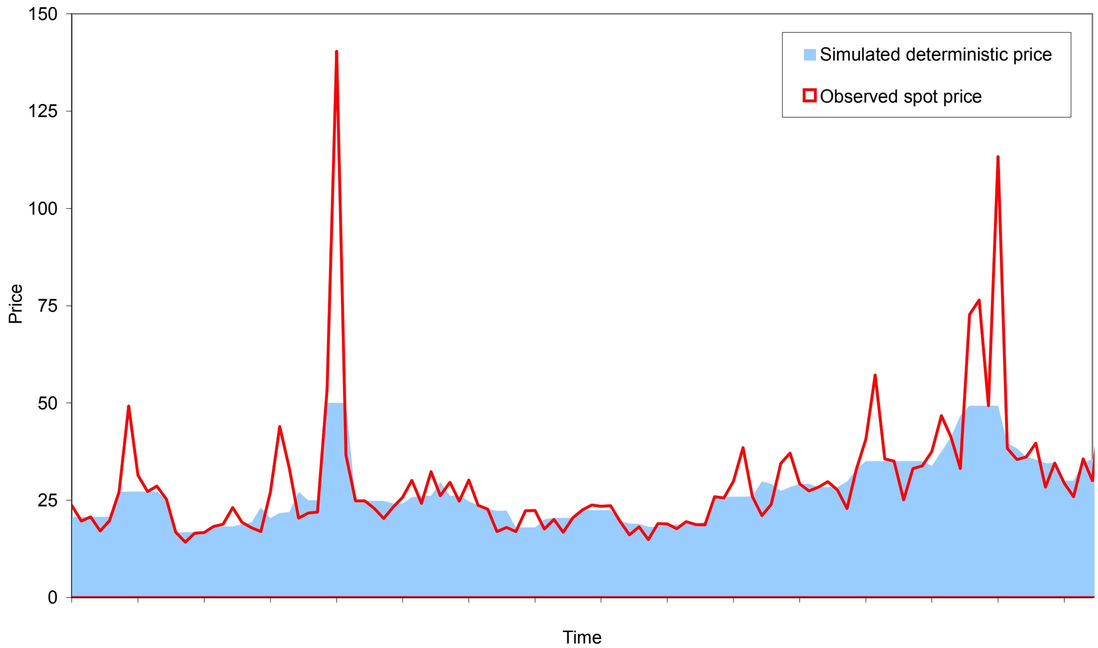

Figure 8.1: Diagram illustrating the results of pre-processing: using a tuned deterministic model to simulate the underlying level of the spot prices.

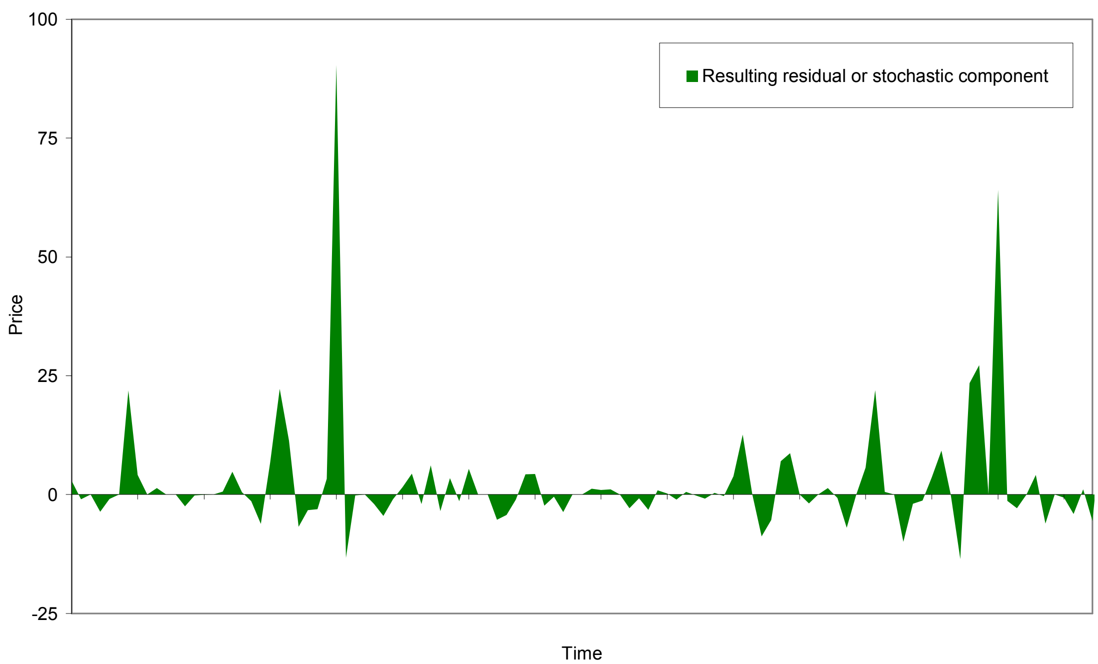

Figure 8.2: Diagram illustrating the discrepancy between actual spot prices and prices simulated with a tuned deterministic model.

8.1.2 Reasons for conducting this research

The overall goal of combining the two types of models is threefold. Firstly, as with all the research presented in this thesis, the aim is to produce better fits to market data (primarily observed spot market prices, but this could also include generation levels). However it should be stressed that calibrating the Cournot model, rather than purely improving the fit of an existing model to historic data, is the major goal of this chapter, hence the majority of the discussion in this chapter revolves around this task.

Secondly, the major reason for calibrating a bottom-up model is so that it can be of use in analysing hypothetical market situations for which market data has not yet been observed. This may be for situations such as after the divestiture of a large generating company, or market entry, to which Cournot models are already applied. In such situations, the use of top-down models on their own is completely inappropriate, as their strength is more in explaining factors such as price volatility that have occurred in the past, given previous market conditions. They are unable to predict the behaviour of market participants and the subsequent behaviour of spot prices if the market were to undergo, for example, a major structural change. Bottom-up models are able to predict this behaviour, but their results are likely to be much more credible if they have been tuned to produce realistic results given current market conditions. Once this has been achieved, assumptions in the model can then be altered to enable prediction of the likely outcomes.

Improving the fit of an existing top-down price model to market prices may well have been achieved through incorporating the physical inputs used by the Cournot model. However, bottom-up models will always find application in market modelling, and calibrating such a model so that it matches observed outcomes in an existing market increases the credibility of such a model for use in future applications.

Finally, linked to the second goal above is the need to produce better tools for testing market designs and forecasting market behaviour. The perfect electricity market design is yet to be developed anywhere in the world, and employing a tool that is able to mimic current market behaviour and forecast future behaviour when testing such designs is likely to give more plausible outcomes.

8.1.3 Structure of this chapter

The research in this chapter is presented in two steps, following the pre-processing and post-processing methods detailed above. The rest of this chapter is structured as follows. The following section gives background on the previous calibration of Cournot models in electricity markets, and details some of the important variables in Cournot models. Section 8.3 contains information on the Cournot model used in this study, and the data included in the calibration. The pre- and post-processing methodology is detailed in Section 8.4, with the results of the study in Section 8.5. Finally, conclusions and discussions regarding the results are provided in Section 8.6.

8.2 Application of Cournot models to electricity markets

As previously stated, Cournot models have been used widely in electricity markets to predict and model the outcomes of hypothetical market situations (see Anderssen and Bergmen, 1995; Arellano, 2002; Neuhoff et al., 2005; among others). They have been extended from simple two-player single-region games, to include extensions such as many players, competitive fringe generation and transmission constraints. However, as Bushnell (2003) notes, “Despite their broad application to electricity markets in the academic literature, oligopoly models have met with substantial scepticism in the policy arena”. He quotes Frame and Joskow (1998), who were “not aware of any significant empirical support for the Cournot model providing accurate predictions of prices in any market, let alone an electricity market.” The purpose of our research is not to extend the literature on the technical side of Cournot modelling, or market modelling in general. Its focus instead is purely on the calibration of one particular Cournot model, and the comparison of the market outcomes predicted by the model with real market outcomes, in order to show that Cournot models can in fact usefully predict market prices.

Cournot models may be deemed by some to be inappropriate for application to electricity markets due to the fact that in reality, market participants offer a set of generation quantities at different prices rather than a single, fixed quantity, independent of price. However, the models are appealing for (among other things) the fact noted by Wolak and Patrick (1997) and echoed by both Borenstein and Bushnell (1999) and Green (2004) that it is possible for generators in oligopolies to increase market prices by reducing their output, which is precisely how Cournot players are predicted to behave.

Using market models to examine observed rather than predicted market behaviour has only been undertaken in the past few years. In a recent study, Evans and Green (2005) apply a Supply Function Equilibria (SFE) model to the British electricity market to simulate prices from 1997 to 2004. Their study involved comparing the relationship between the simulated prices and actual prices across time, to assess whether or not market decentralisation had led to a reduction in prices. While the use of SFE and other market models in an electricity market setting has been researched extensively, the application of non-Cournot models is outside the scope of this study. Green (2004) asked whether or not British generators played Cournot games, however his approach was to look at capacity withholding instead of analysing the actual market prices that resulted. His approach followed the logic that generators playing Cournot games and competing on quantity will withhold more capacity than if they were perfectly competitive. This reduction in capacity is in effect what drives prices to be higher than perfectly competitive levels.

The most widely-referenced study involving the use of a Cournot model to analyse historic electricity prices was completed by Bushnell (2003). He used the Cournot model of Borenstein and Bushnell (1999) to simulate prices in California during the summer of 2000, and showed that the mean monthly prices that actually occurred were much closer to his mean simulated Cournot prices than his mean simulated perfectly competitive prices. In fact, the mean price in September 2000 was slightly *greater* than his mean Cournot price for that month. He then simulated the effects of reducing market concentration and introducing demand elasticity (among other adjustments) to analyse how prices in 2000 could have been reduced. In the context of our work, the major point of note from Bushnell's work is that Cournot models can, despite their apparent inappropriateness for use in electricity markets, make accurate predictions of what electricity prices will be.

In a more recent study, Neuhoff et al. (2005) compare the prices modelled by four network-constrained Cournot models of the electricity network encompassing Belgium, France, Germany and the Netherlands. They model demand at each node with a linear demand function with elasticity -0.1, but do not include forward contracts in their models. The aim of their study is to examine the effects on the modelled prices of the various assumptions of the four models, such as whether or not the Cournot players are able to anticipate the effects of their decisions on fringe behaviour and transmission prices. They noted that how such assumptions are made does have a substantial impact on the level of simulated prices.

The majority of the prices simulated by the models in their study overestimate prices, however they note that matching observed market behaviour was *not* one of their aims. They state that it is tempting to reduce the level of simulated prices, which could be accomplished by three separate methods: firstly, increasing the elasticity of demand so that generating companies have less opportunity to exercise market power; secondly, modelling forward contracts in the market; and thirdly, assuming that one or more Cournot players act instead as perfect competitors, under the (implied or explicit) threat of regulation. They refer to the results of a survey undertaken at a workshop attended by "eleven experts in power market modelling and regulation", who, interestingly, placed very little importance on such models actually matching market behaviour, as opposed to the insights provided by modelling. This seems extraordinary; without knowing whether or not models match reality to start with, how can any insights regarding hypothetical market situations be given any credibility?

8.3 Australian price data and the Cournot model

The market and time period we modelled in this study was Australia's National Electricity Market (NEM) from 1 January 2003 – 31 December 2004, for a total of 731 days. Over this time period, the NEM comprised four regions: Queensland (QLD), New South Wales (NSW), Victoria (VIC) and South Australia (SA). The four regions are linked together through interconnectors between each of the adjacent regions, however transmission capacity across these interconnectors is limited and at times when transmission is constrained, prices can vary markedly between regions. Approximately 97% of the generation capacity in the NEM is thermal generation, making it ideally suited to a bottom-up model as estimating marginal costs is quite straightforward compared to a market dominated by hydro generation.

8.3.1 The price series

The majority of previous studies on price modelling mentioned in this thesis focus on daily average price data, however aggregating data from across 24 hours into a single point for each day is too crude an approximation for use in a Cournot model. In the NEM, prices are set at five-minute intervals, and loads can fluctuate wildly depending on the time of day and the time of the year. Winter loads tend to follow the same patterns as in New Zealand, with peaks in the morning and evening. However, in Australia the most extreme load fluctuations are experienced in the summer, when occasional heat waves during weekday afternoons force people to turn on their air-conditioning at times when loads are already at their peak. Often the extreme loads that result can place extreme pressure on the security of supply, and force prices to the Value of Lost Load (AUS$10,000/MWh). As a result of these extreme intra-day load fluctuations, we decided it was necessary to use data in a higher frequency than daily averages, and, using the method detailed below, we separated the data from each day into three distinct periods: "peak", "shoulder" and "off-peak".

In determining which observations in each day should be allocated to each period, it is not enough to choose specific time period(s) for each day of the year, for example 8am-9am and 5pm-7pm, to be the "peak" period, for the reasons mentioned above. The method we used to determine the allocation of observations to periods was therefore based on the total system load (i.e. the sum of the loads from each of the four regions) observed during each day. Theoretically, the periods with the highest total system loads should correspond to the periods with the highest spot prices, although in reality this is will not always be the case as peak loads do not always occur at the same time in each region.

We decided in this study that the number of observations per day to allocate to each period would be constant across the two years. In reality, the number of peak half-hours per day et cetera depends on factors such as the season; however, a detailed examination of load patterns across time was outside the scope of this study and is left for future research. In order to assess how many half-hourly observations should be allocated to each of the three periods, within each of the 731 days we ranked the total system loads from each half hour from 1 to 48. Then, for each half-hour, we calculated the load as a percentage of the maximum daily load during that day. Finally, we sorted the loads by rank, and found the average percentage of maximum load for each rank (1 to 48) across the 731 days, giving the graph of these average percentages versus the rankings shown in Figure 8.3. This graph shows, for example, that the 25th largest total load per day during the sample was, on average, 90% of the largest load that day.

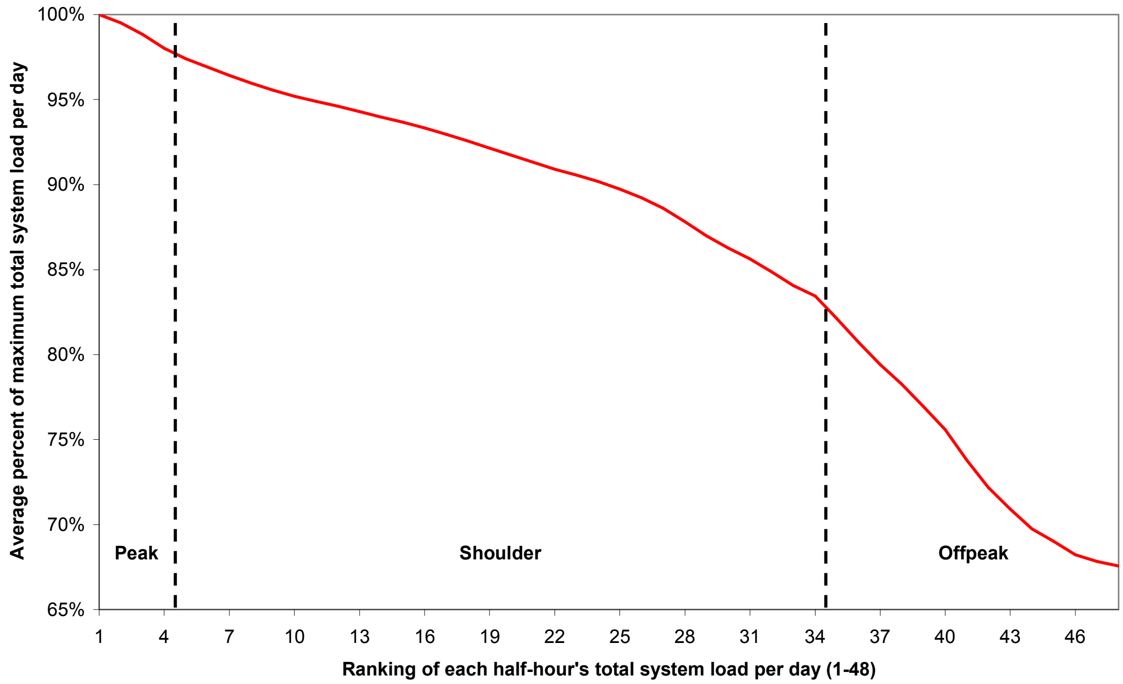

Figure 8.3: Percent of maximum NEM system load versus rank of total NEM system load for each half-hour, sorted by day, January 2003 – December 2004

As mentioned earlier, we aimed to allocate a constant number of half-hourly periods per day to each of the peak, shoulder and off-peak periods, with the major decision being how many periods per day to allocate to each. By observing relationship between the two variables in Figure 8.3, and through our experience with Australian load conditions, we decided to allocate the four half-hourly periods with the greatest total system loads per day to the peak period, the next 30 greatest half-hourly load periods to the shoulder period and the remaining fourteen periods to the off-peak period. While our method of allocation was arbitrary in the context of this exploratory analysis, no doubt a more formal method could have been used and would be appropriate in a more thorough study.

It would appear logical that for a given series of peak, shoulder or off-peak observations, the Cournot model should be solved for each of the 731 days in the study to form an underlying path for the actual prices. The behaviour of these prices would then have been driven largely by the loads observed on those days. However, had it been decided at the start of 2003 that two years of price forecasts were required, there should be no reason why the underlying price forecast should be any different on (say) a Wednesday than it would be the preceding day on Tuesday. While there should be notable differences on Saturdays and Sundays (and public holidays), differences during the week would be harder to forecast. For this reason the final set of observed loads was reduced to three observations per week: Saturday load, Sunday load, and weekday load, which was the average of the five weekday loads from Monday to Friday. Plotted as time series, the three load series (in terms of total system load) used as input to the Cournot model are shown in Figure 8.4.

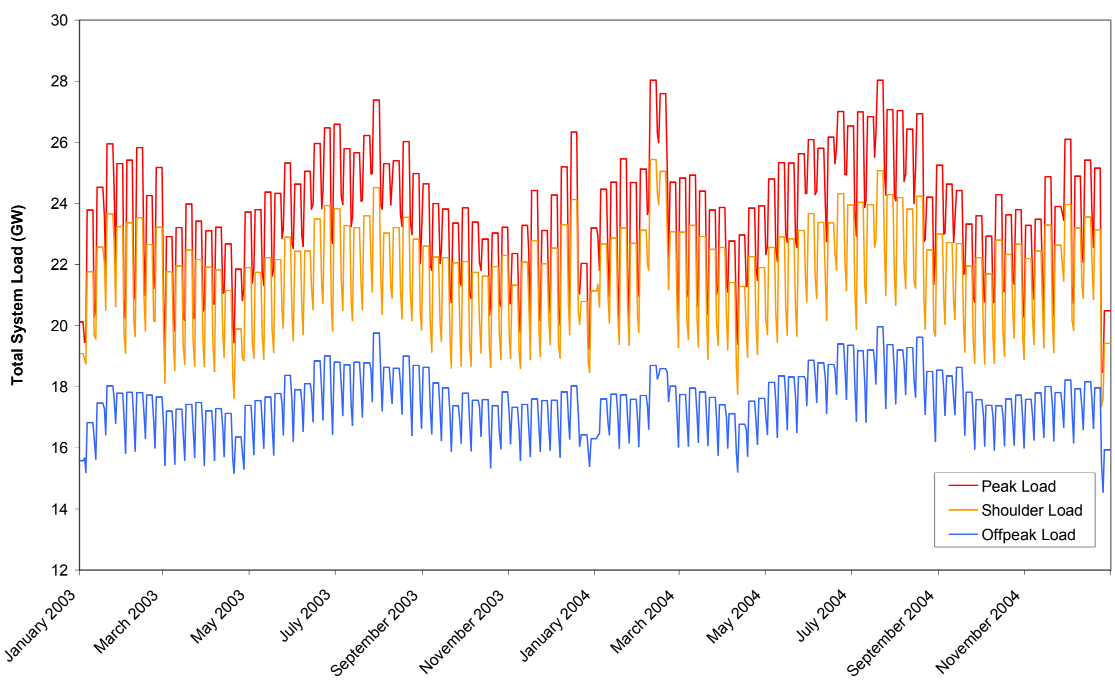

Figure 8.4: Total system load used as input for the Cournot model – peak, shoulder and off-peak series

Unlike each of the load series, the price series, however, were not averaged across the week, in order to preserve some of the variation that could be explained later in the post-processing stage of the study. However, prices were averaged within each load period in each day. Instead of using the resulting prices for each state in the NEM, we also decided to focus simply on the VIC prices for the post-processing stage of the study. The three price series for VIC are shown in Figure 8.5. Note that it is possible for the price in the shoulder series on a certain day to be higher than the corresponding price in the peak series. This is because the observations were aggregated based on the total system load, and not the individual region loads nor the observed prices. The periods with the highest total system loads may not necessarily correspond to the periods with the highest loads in each of the four regions. Also, due to transmission constraints, random outages and other unexpected events, the highest prices during the day are not necessarily caused by the highest loads.

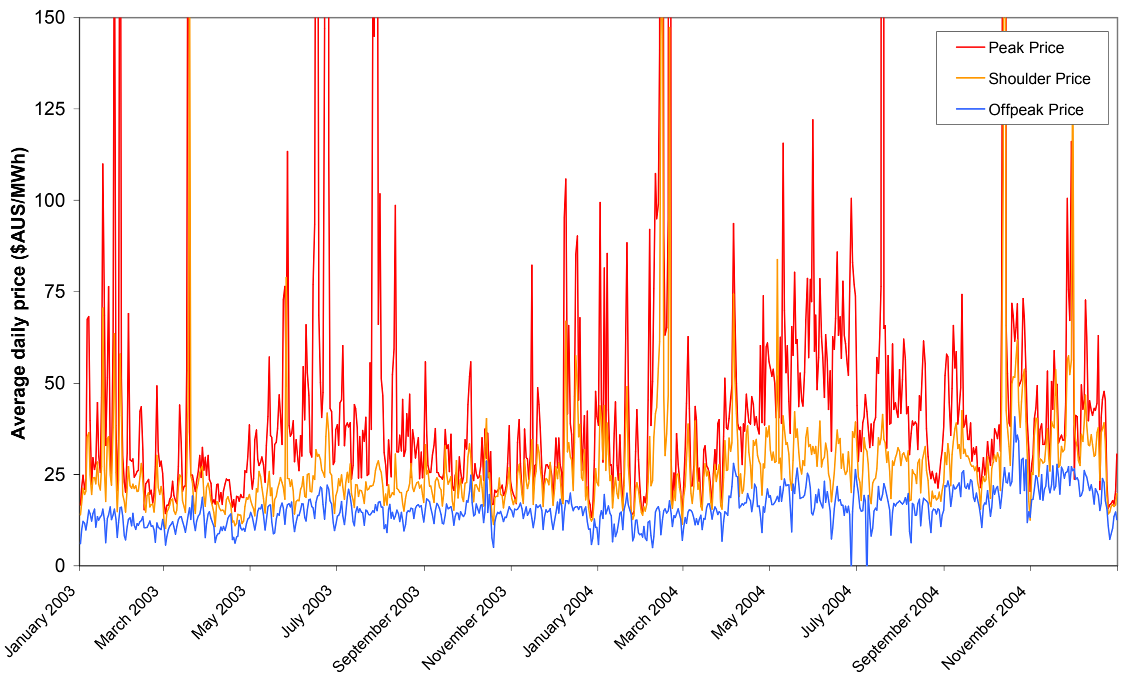

Figure 8.5: Average daily spot price in Victoria for each of the three series – peak, shoulder and off-peak, 1 January 2003 – 31 December 2004

.

As an initial step, each of these three series was fitted using the Escribano, Peña and Villaplana (EPV; 2002) model detailed in Chapter 3. The estimates of each parameter are listed in Appendix I, with the interesting features of these results detailed later in this chapter, in Section 8.5.1. At this stage, note that the highest prices occurred during summer and winter over the two years, and price volatility was considerably less in autumn and spring. As one would expect, the peak series was the most volatile with the off-peak series being the least.

8.3.2 The elasticity of demand and the contract level

Two of the crucial inputs into any Cournot model (or any model of an electricity market) are the elasticity of market demand and the contract level, but neither is publicly available in the same way as load and generator data.

8.3.2.1 The elasticity of demand

In market models, the elasticity of demand refers to the rate at which consumers will change their levels of consumption given an increase in prices. The first premise associated with the elasticity of "normal" goods is that an increase in prices will result in a decrease in consumption; therefore, the elasticity of demand will be negative. In the long term (or long run), consumers have the ability to change their existing technology and infrastructure (such as switching from electric to gas heating) in order to reduce consumption and minimize the effects of a price increase; however they do not have this ability in the short run.

The second premise follows from this – demand is more elastic (i.e. more responsive to changes in the price) in the long run than in the short run. Demand for electricity in the short run is almost completely inelastic, due to the fact that many consumers (such as individual households) do not pay the spot price for electricity, instead paying a fixed rate determined by their retailer. However, in the long run, any changes in the overall level of spot prices filter through to the prices set by the retailers. Therefore, consumers' behaviour may change given a long enough time period.

8.3.2.2 The contract level

As with many financial markets, most electricity markets provide consumers (including retailers) with the ability to purchase contracts for electricity. These contracts give the buyer (usually a large consumer or retailer) the option (but not the obligation) to purchase from the seller (usually a generating company) a specified amount of electricity *K*, for a specified cost per unit *PK* (the 'strike price'), at a specified time. This gives consumers the ability to hedge against high spot prices, meanwhile giving generating companies a guaranteed stream of income regardless of the actual outturn spot prices. The contract is not free, and its value (and price) depends on the 'strike price', the volatility of the underlying spot price being hedged against and the time until the contract can be exercised. In contrast to other markets, in electricity markets there is no way of guaranteeing that electricity produced by one generating company will be consumed by any specific consumer, and market-clearing models ensure that demand always equals supply. As a result, there is no obligation on the seller of a contract to generate *K* units of electricity, nor on the buyer to consume any electricity. The contracts just guarantee a fixed price paid and received for any electricity consumed, up to the contract amount *K*.

One of the more common electricity contracts currently used in many markets around the world, including the NEM, is a two-way contract for differences (CFD). These have the effect of "locking in" the price of electricity paid by consumers and received by producers. As is the case without contracts, during the period to which the option applies, the generating company receives the spot price *p* for each MW of power it produces, and the consumer pays the spot price *p* for each MW of power it uses. However, after the contract has expired, the two-way CFD requires that if the spot price *p* was greater than the strike price *PK*, the seller of the CFD has to refund the buyer the difference between *p* and *PK* for each of the *K* units, regardless of how many the buyer actually used. As the contract is two-way, if the spot price *p* was less than the strike price *PK*, the buyer must then pay the difference to the seller for each of the *K* units. In this way, the two-way CFD requires merely an *ex post* financial transaction, and does not influence the consumption decisions of the buyer of the contract[2]. It is in both parties' interests to buy and sell contracts – generators can decrease the risk they face from spot prices being low (while guaranteeing income), and consumers can reduce the risk they face from high spot prices.

[2] Under the assumption that the consumption decisions of the buyer of the contract do not influence the spot price, the *ex post* financial transaction (and the cost of the CFD) can be treated as a sunk cost (see Batstone, 2003).

By locking in the price that electricity is traded for, the risk associated with variable profit margins is reduced for both parties.

Contract levels are particularly important in solving a market model, as generators' behaviour will be very different depending on the extent to which they are contracted. In Cournot models, generators exercise market power by reducing their capacity in order to increase the spot price. However, if generators have a guaranteed stream of income regardless of the spot price, and in fact will have to refund their customers via a CFD if the spot price is greater than the strike price in the contract, they do not have this incentive to raise prices. The profit-maximising strategy of a contracted generator is to offer the amount of generation for which it is contracted at SRMC, which is presumably lower than the contract strike price. This ensures that, if they are dispatched, they will receive (spot price – SRMC) for each unit that generates, which will be greater than any amount (spot price – PK) they have to repay customers via a CFD. If they offer at cost and are not dispatched, the spot price will be less than the strike price, guaranteeing them an *ex post* payment from their contract customers anyway. Generators still have the opportunity to withhold any excess generation they may have over and above their contract level, to drive up the spot price and maximise their profit over the whole portfolio, but the amount of generation for which they have an incentive to withhold is much less than if they were not contracted. If they are contracted for 100% of their capacity, their profit-maximising strategy will be to offer all their output at cost, effectively forcing them to act as perfect competitors.

It turns out that if generators are over-contracted (i.e. contracted for an amount greater than they are able to supply), it is actually in their interests to *lower* the price rather than raise it, even to levels below the marginal cost of their most expensive unit dispatched. In the *ex post* settlement they will receive more for each MW they are contracted for, the lower the spot price is[3]. Hence the extent to which generators are contracted greatly influences market outcomes.

8.3.2.3 These parameters in the Cournot context

As mentioned earlier, neither the elasticity of demand nor the contract levels are publicly available. Neither is well determined from first principles, as it is likely that generating companies do not assign specific values to either parameter before determining their optimal actions. It is therefore logical to attempt to infer from market data the values of these parameters participants may be using.

In fact, it is not even clear how either parameter should be defined in the context of Cournot modelling. In a Cournot model, the elasticity of market demand theoretically refers to the elasticity in the short run only. Since an electricity market is not actually Cournot, and participants in the market are unlikely to act as Cournot players, the "elasticity of demand" somehow also serves as a proxy for the response of other generators to changes in price. Each Cournot competitor picks their BR output to offer into the market based on the demand curve faced by the market and the likely actions of other players in the market. Therefore, they have to choose their strategy based on the residual demand they face – that is, that portion of demand not met by their competitors. If the market is highly competitive, with many participants each having similar marginal costs of generation, then if a generator offers their output at too high a price their share of the market is likely to be taken by a competitor. The residual demand they face in this case is said to be elastic – a small change in price will lead to a large change in the amount of generation offered. However, if competition is not tight, each generator has a greater ability to offer their output at a higher price and still be dispatched. Residual demand in this case is inelastic – a small change in price leads to a small change in the amount of generation offered. Therefore, the elasticity of market demand in a Cournot

[3] See Wolak (2000) and Batstone (2003) for a more thorough explanation of the influence of contracts on generator bidding behaviour and spot prices.

model represents not only the response of consumers to a change in price, but also the response of generators.

As Wolak (2003a) illustrates, the elasticity of the residual demand curve faced by a generator is directly linked to the Lerner Index[4], which measures that generator's ability to exercise market power (i.e. to raise the spot price above its short-run marginal cost of generation). If residual demand is elastic, then a generator cannot raise the price without losing market share or decreasing consumers' demand. However, in a market situation with inelastic residual demand, a generator can still raise the price without eliciting too great a response from either its competitors or consumers.

There is an obvious link between the elasticity of demand and the contract level. As explained, the more inelastic the market demand, the higher generators can raise the price without provoking a reduction in demand. However, if they are heavily contracted, their market power is diminished and they do not have the same incentive to raise prices. In some cases, as alluded to earlier, a high level of contracting will create an incentive to decrease spot prices below marginal costs and increase (rather than withhold) generation[5]. If demand is particularly inelastic, consumers are less able to reduce their consumption when prices increase, and they have an incentive to purchase contracts to reduce the risk of being exposed to high prices. If the market is particularly competitive and residual demand is elastic, then it is in generators' interests to sell contracts and increase generation to guarantee income. Therefore, when demand is both elastic and inelastic, contracting is desirable.

Just as the elasticity of demand is not well defined in the Cournot context, "contract levels" may also serve as a proxy for other price-restraining factors such as regulation, or the threat thereof. As mentioned earlier, a heavily-contracted firm will price its generation near (or even below) marginal cost and increase its output. The same

[4] As explained in Chapter 2, the Lerner Index is one of the most common measures of market power.

[5] See Wolak (2000).

behaviour would be observed if a firm was facing the threat of regulation. Regulators of electricity markets pay particular attention to situations when spot prices are above marginal generating costs, and in order to reduce this possibly unwanted attention, generators are likely to keep output levels high and prices low.

In general, the elasticity of demand can only be estimated from observed market data, while contract levels are confidential, and, for vertically-integrated firms, often only implicit in the need to meet retail load requirements. Nor is it clear over what time horizon participants will want to assess either parameter[6]. Thus it seems reasonable, and the only available option, to deduce the values of both these parameters from observed market behaviour. As the elasticity of demand in a Cournot model is likely to represent more than just the response of consumers to changes in the spot price, and the Cournot contract level is similarly not well defined, we label the two parameters inputted to the Cournot model as the Pseudo Elasticity of Demand (PED) and the Pseudo Contract Level (PCL). Throughout the remainder of this chapter, when these input parameters are referred to (as opposed to the more general economic concepts) the labels PED and PCL will be used.

8.3.3 The Cournot model

The Cournot model used in this study was created by CRA International (Asia-Pacific) Ltd, and is described in Chattopadhyay (2004). It is named T-CONE, an acronym for the Transmission-constrained Cournot-Nash Equilibrium model. As documentation for the model states, "T-CONE formulates this problem as what is known as a mixed-

[6] An example of the ambiguity of the time horizon is the fact that generators' *behaviour* with respect to the amount they are contracted may change over time, even if their actual level of contracts does not change. If a generator is nearing the end of a fixed-term contract with a relatively low strike price, it may be in their interests to act as though they are not contracted and raise the spot price to ensure that the strike price of their next fixed-term contracts is greater. Green (2003) echoes similar sentiments when he asks, "Does long-term contracting, or its logical extension of vertical integration, provide sufficient incentives to keep prices down, or would generators wish to raise wholesale prices in order to raise retail prices and the price of future contracts?"

complementarity problem. This is due to the fact that finding a Cournot-Nash equilibrium involves simultaneously solving for values of the price and quantity variables; standard optimisation models hold one or other of these variables fixed" (CRA International (Asia-Pacific) Ltd, 2003).

The input data for T-CONE includes:

- Cost: the short run marginal cost (SRMC) of each generating plant.
- Load: an estimate of the demand curve in each period for each region.
- Generation: available generating capacity of each unit in each period, the location of that capacity and the generating company it is owned by.
- Transmission capacity: the links and transmission capacity between the different regions in the market.
- PCL: an estimate of the CFD quantity for each generating company.
- PED: an estimate of the elasticity of demand for each region.

In this study, all the inputs listed above are held constant throughout the sample period, aside from the load (described in the previous section), the PCL, and the PED. Stochasticity in the inputs (for example stochastic fuel prices or plant outages), could have been layered in through a Monte Carlo simulation framework. This should be considered in any future study, however running Monte Carlo bottom-up models is a very computer- and time-intensive exercise and single-run simulations served the purpose of this exploratory study.

With only load varying from day to day in each simulation, the Cournot price in each region is actually calculated only as a complicated function of load, including the PCL and PED. If forecasting prices were the only aim of this chapter, then it would have made more sense to calibrate a simpler function of load for the price in each region, however this function would have been complicated by the interaction between prices in each of the regions and the transmission constraints. As described earlier in this chapter, this would also not have achieved the aim of calibrating the Cournot model.

T-CONE is a static equilibrium model, solving each period independently. This means that complexities such as hydro storage over time are not modelled in our study. Also, "Neither short-term (e.g. transmission, ancillary services, unit operating characteristics) or long-term (new entry) system dynamics are modelled due to the increased difficulty of finding the equilibrium solution."

T-CONE is solved as a quadratic program, and the transmission constraints within and between each region are imposed directly on the program itself. The quadratic programming model maximises the "welfare-adjusted total market benefit" (Chattopadhyay, 2004), which is the sum of the market benefit at the perfect competition solution less the deadweight loss associated with each of the Cournot players having some element of market power. The solution to this problem produces the maximum total profit of all the generators; this solution "automatically ensures that individual genco profit is maximized and it is a Cournot-Nash equilibrium". Shadow (or nodal) prices are found for each region on the nodal balance constraints for each region. These constraints ensure that generation injected into a node, plus the flow into that node from other nodes, minus the flows from that node to other nodes, equals the demand for that node. This method of price calculation is consistent with market-clearing software, and it is these shadow prices that are reported as the regional prices in this chapter.

Further details of the model, as well as the full specification of the quadratic program and discussion around the treatment of transmission constraints, can be found in Chattopadhyay (2004).

In its practical operating state in the GAMS platform, T-CONE solves in two steps for each individual day. First, the perfectly competitive solution is found, which gives the least-cost dispatch and price for the given market demand curve. Second, the deadweight loss term is added to the objective function of the model, which solves to find the Cournot-Nash Equilibrium price and dispatch amount. The market demand curve is the same for both steps of the model, and is determined (as described in the following paragraph) by the load and price pairing, and the PED. However, instead of being a fixed MW amount, the contract level for each generating company is calculated as a percentage of that company's total least-cost dispatch, which in turn is dependant on the elasticity. Therefore the absolute contract level (in MW terms) for each solution of the model actually varies with every combination of PED and contract percentage level (so the PCL is expressed as a percentage). The PCL for each company is a required input for T-CONE.

The demand curve used by T-CONE is a linear inverse demand of the form $D(P) = A - bP$, where P is the price. A more complete study would have used both linear and constant elasticity demand curves; however the use of a different form of demand curve is left for further research. Criticisms of linear demand curves in studies including Borenstein and Bushnell (1999), such as the fact that they can intercept the price axis at an unrealistically low point[7], should, however, be noted. As with the constant elasticity curves in their study, each demand curve has its slope determined by the PED, however the position of the curve on the price-generation axes will vary. In general, the curve must be calculated to pass through a certain load-price reference point, which then determines the position of the curve on these axes.

While the focus of the calibration in this study is on the elasticity of end-use demand, several studies of Cournot models in electricity markets focus on the residual demand elasticity faced by a set of Cournot firms. As Bushnell (personal communication, 2007) states:

> "Much of this residual elasticity is coming from the supply of non-Cournot (fringe) firms. This can be estimated from basic firm structure (i.e.

---

[7] A downward-sloping linear demand curve, by definition, will intercept the price axis (i.e. quantity demanded = 0) at some point. However, these authors note that in their study, an elasticity of just -0.1 resulted in the demand curve intercepting the price axis at US$1023/MWh. They state that, "It is unrealistic to assume demand would be completely curtailed by prices in that range", which is fair, as prices in the NEM also reach far greater prices with little significant impact on demand.

> installed capacity) as in Borenstein and Bushnell (2000), or more recently, actual market data have been used to estimate a fringe supply elasticity, which is then applied to a residual demand for the strategic firms. Bushnell, Mansur and Saravia (first version 2005) use this approach, as does Puller (2007), to fit Cournot models to market data. … One implication of using a "fringe supply" approach to deriving residual demand elasticity is that the linear functional form is not so unsatisfactory. … It depends on whether the capacity constraints of the fringe firms are likely to be binding, providing a sharp curvature to the residual demand curve.[8]"

In our study, each of the four regions has its own demand curve for each day in the sample, the position of which is determined by the load that day in that region and a reference price. To assign each of the four regions a reference price for each of the 731 days in the sample, we took a list of all the loads and all the prices for each region over the sample period, and sorted both lists in order from least to greatest. This meant that the greatest load was paired with the greatest price; the 100th greatest load was paired with the 100th greatest price; and so on. The individual regions' pairings then became our reference points for the demand curves in T-CONE. These reference points for the four regions' demand curves are shown Figure 8.6, and show clearly that the reference price increases as the load increases. One can imagine that each of these points has a downward-sloping linear demand curve drawn through it.

---

[8] Bushnell also notes that one advantage of using a residual elasticity estimated from market data is that it effectively fixes one of the two variables being tuned in this chapter, the elasticity of demand, allowing the potential values of the contract level to be examined more thoroughly.

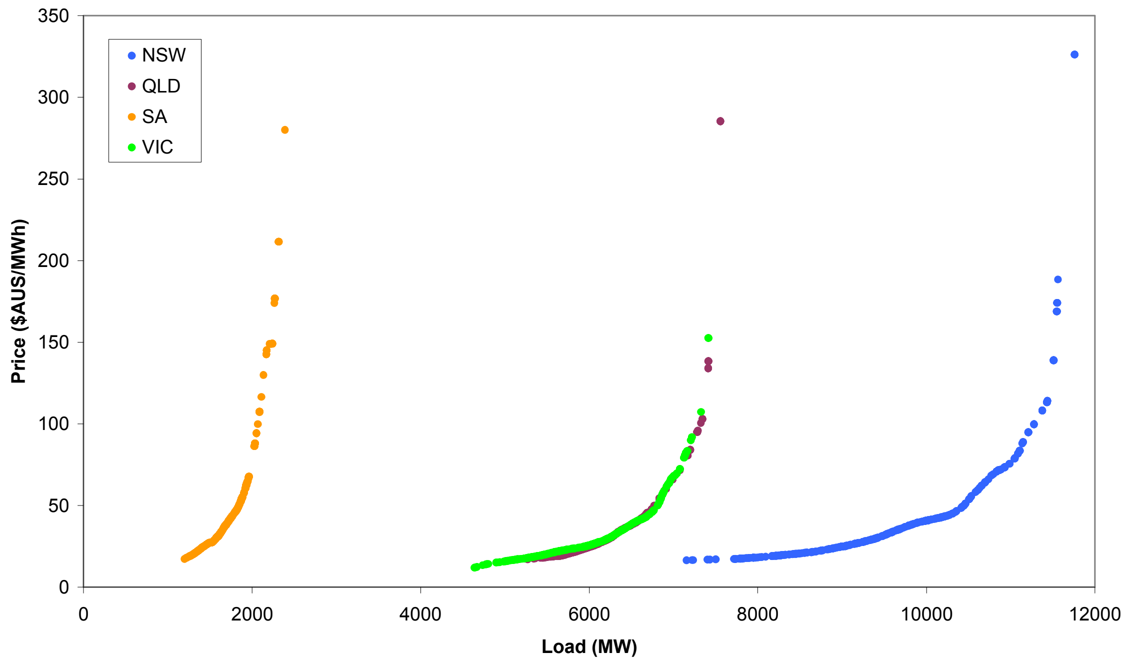

Figure 8.6: Load and price pairings for each region over the sample period, 1 January 2003 – 31 December 2004.

Aside from the PCL and PED, which are the inputs into the model that we defined, all data on generation costs and capacities, prices, load and transmission were sourced from CRA International's NEM database. The T-CONE model was treated as a black box in this study, in that inputs were given to the model and outputs received, without any further investigation into how the inputs were converted into outputs. Because of this, a thorough and exhaustive test of the results of varying the PCL and PED was completed, to verify that T-CONE produces appropriate results. The results of the verification are in Appendix H.

8.4 Methodology

As mentioned above, there are two steps in this research – pre-processing, and post-processing. The pre-processing step in this case involves determining the single combination of the PED and PCL that produces the most accurate fit of the underlying level of actual price time series. The T-CONE model also calculates generation levels for each plant in the NEM, as well as flows between the regions via the various interconnectors. Each of these could also be compared with actual levels to test the model's performance. However, in this study we chose to examine only the fit of the prices provided by T-CONE, leaving open the option of examining generation and flow modelling accuracy for future research.

In this exploratory analysis, we chose to keep both the PCL and PED constant across all companies and all regions respectively, and across every observation in the two years of the sample. These parameters no doubt vary widely across time, companies and regions, and are a major cause of price volatility, along with other factors such as plant and transmission outages (both planned and unplanned)[9]. However we ruled the effects on prices of varying these parameters outside the scope of this study. Therefore for each of the 731 days in the sample period, T-CONE had the same available generating and transmission capacity, the same PED and the same PCL.

We did, however, vary the PCL and the PED across each of the peak, shoulder and off-peak periods. In the NEM, forward contracts often cover specific periods of the day, and due to the differing load throughout the day and different consumer requirements for electricity, the short-term elasticity of demand is also likely to fluctuate depending on the time of day[10]. For example, if everyone were exposed to spot market price fluctuations, businesses that require electricity to operate throughout the period from 8am to 5pm may have a much less elastic individual demand curve than individual people or families do when they are at home outside of those hours. In summer, when peak loads are higher, a

---

[9] The contract level and elasticity of demand in electricity markets no doubt vary throughout the year, and the amount of available generating capacity is not constant throughout. Prices are likely to be higher, and much more volatile, when, for example, a large base load generator is offline for maintenance. However, as mentioned in Section 8.3, from a long-range price-forecasting point of view it may not be possible to know exactly when such changes will occur. As mentioned in Section 8.1.1, if the aim of the forecasts was to forecast the overall price distribution rather than a specific price time series, then parameters such as the PED and PCL could be varied randomly in a Monte Carlo simulation setting.

[10] For example, Bushnell and Mansur (2005) found that during a period of electricity price increases in San Diego, reductions in load were greater during peak afternoon hours than at other times of the day.

person may be much less likely to consider the price of electricity when they turn on their air conditioning at work to cool down during the day than they may be when choosing what to cook for their evening meal or how to heat their house in winter.

After further consideration of consumer behaviour, the peak series was split into weekday peak and weekend (and public holiday) peak periods, but only for the purposes of calibrating the Cournot model. This is due to the very different conditions influencing peak loads on weekdays compared with weekends. As the shoulder periods contained a much larger number of observations per day it was decided not to split this in the same way, and off-peak behaviour was considered unlikely to vary greatly depending on the day of the week either. The shoulder and off-peak periods should be split, however, in a more complete study.

In order to assess how accurately a particular simulated price series fits the actual price series, and compare between the fits of several different simulated series, we required a measure of goodness of fit. The most commonly used method in the literature for forecasting purposes is the root mean squared error (RMSE). This is calculated by squaring each of the differences (or errors) between the fitted observations and the actual observations, finding the mean squared difference and then taking the square root of that mean squared error. Another criterion making use of the concept of squared errors is the r-squared statistic, which involves calculating the sum of the squared errors and the sum of the squared differences between each observation and the mean observation.

As argued by de Lange, Schavemaker and van der Sluis (2002) among others, these criteria, and others that involve taking arithmetic averages such as the mean absolute error, and the mean absolute percentage error, are not entirely suitable for assessing the goodness of fit for time series that involve extreme outliers, such as high-frequency electricity price time series. Electricity price series contain spikes, the effects of which tend to dominate these traditional fitting criteria. For example, several spikes in the observed price series occurred for reasons which could not be explained by the load or generation data available, and may have been caused by unforeseen and unrecorded transmission outages. As mentioned above, such outages were not taken into account in this study and therefore the price forecasted by T-CONE could not possibly be close to the actual price that results. As these spikes were very large, each of the aforementioned goodness of fit statistics was strongly influenced by these errors, and the difference between the fitting statistics for different simulations was negligible. Even if the whole simulated series were a perfect fit except for the spikes, this would still yield a much lower assessed fit than a simulated series that fitted the spikes well and the rest of the series relatively poorly. As the aim of pre-processing is to model well the underlying level of the price series and leave the spikes and other volatility to the stochastic process (except for the volatility that could be modelled by the Cournot model), the most important thing is that the majority of the non-spike prices are fitted accurately.

In order to measure this fit, it is therefore necessary that the median errors are measured, rather than the mean errors. This is not a new concept in electricity price forecasting, nor in other areas of forecasting, and de Lange et al used the *median* absolute deviation (MAD) rather than the mean absolute deviation in order to filter the effects of outliers from their analysis of electricity spot prices. The MAD is calculated in essentially the same way as the mean absolute deviation, in that the absolute value of the forecasting error for each observation is calculated. Then, instead of calculating the arithmetic average of these absolute errors, the MAD statistic is the median absolute error. The two sample statistics will be similar if there are few outlying observations and the distribution of the underlying variable is symmetric, but they will be dissimilar if there are many observations which cannot be well modelled.

The T-CONE model calculates four different prices for each solution, one for each of the regions in the NEM. The aim of the pre-processing step in this study is to forecast with accuracy the price in each of the four regions, therefore the measure of goodness of fit has to take into account the forecasting error in each region. The statistic chosen for comparison between different simulations is the average of the four regions' MADs. If the median MAD of the four regions had been chosen instead, it is possible that the fit in two of the four regions could have been mistakenly ignored entirely.

Each particular simulation of 731 days provides one overall average MAD, and the relative forecasting performance of each simulation is compared using these statistics. For each of the three series being simulated, a two-dimensional grid of PCLs (from 0% to 110%) and PEDs (from -0.01 to -2) is formed. The contract percentages used for this study are 0%, 25%, 50%, 75%, 90%, 100% and 110%. This range of percentages is used based purely on experience, as no information is available on the extent to which generators in the NEM are contracted.

In contrast, studies exist on both the short- and long-run elasticity of demand in electricity markets including the NEM (see Langmore & Dufty (2004) and National Institute of Economic and Industry Research (NIEIR; 2005)). But the elasticities presented range widely, from -0.16 to -0.7. Bushnell introduces a constant elasticity of -0.075 in his 2003 study, which he says is more inelastic than other estimates; his justification is that most other estimates did not come from as high-frequency price data as he was using (hourly). He also uses values of -0.05, -0.10, and -0.20 in his 2002 Cournot model of electricity markets in the western United States. As a result of this range, and the fact that three different load situations are being analysed in this study, the PEDs selected are -0.01, -0.05, -0.15, -0.3, -0.5, -1, -1.5 and -2. To illustrate the meaning of these values, an elasticity of demand of -0.15 means that if electricity prices were to rise by 10%, a 1.5% fall in demand for electricity would result. Therefore, an elasticity of -0.01 suggests that demand is highly inelastic (i.e. electricity consumers' demand changes very little with changes in the spot price), and an elasticity of -2 suggests that demand is highly elastic.

The eight elasticities and seven contract percentages present a grid of 56 different combinations of input values, and from each of the resulting simulations a mean MAD can be calculated. The combination that provides the lowest MAD is deemed the most appropriate for determining the deterministic price level across each of the four regions.

Once the appropriate combination of parameters for each load period has been selected, the pre-processing step is complete. Then, for each observed series of prices (although in this study we focus only on Victorian prices), the residual or error price series (i.e. the component of the price series not explained by the Cournot model) is calculated by subtracting each observation in the simulated series from each observation in the actual series. The resulting series of 731 residuals is then fitted using the EPV stochastic process model (equation [5] in Chapter 3). Once this fitting is complete, assessments can be made as to whether or not the Cournot model provides a superior fit to the deterministic component of the EPV model.

The whole methodology is presented diagrammatically in Figure 8.7 below. The pre-processing inputs, the PED and PCL are chosen by hand, and entered into T-CONE, which is treated as a black box. The combination of the PED and PCL which yields the best fit to the underlying level of the observed prices, based on the MAD criterion, is then used to produce the simulated deterministic component of the observed prices (the blue box). The residual variation in the spot prices (the red box minus the blue box) is then used as input into the top-down stochastic process model, and the parameters of the model are estimated using the CML estimation procedure.

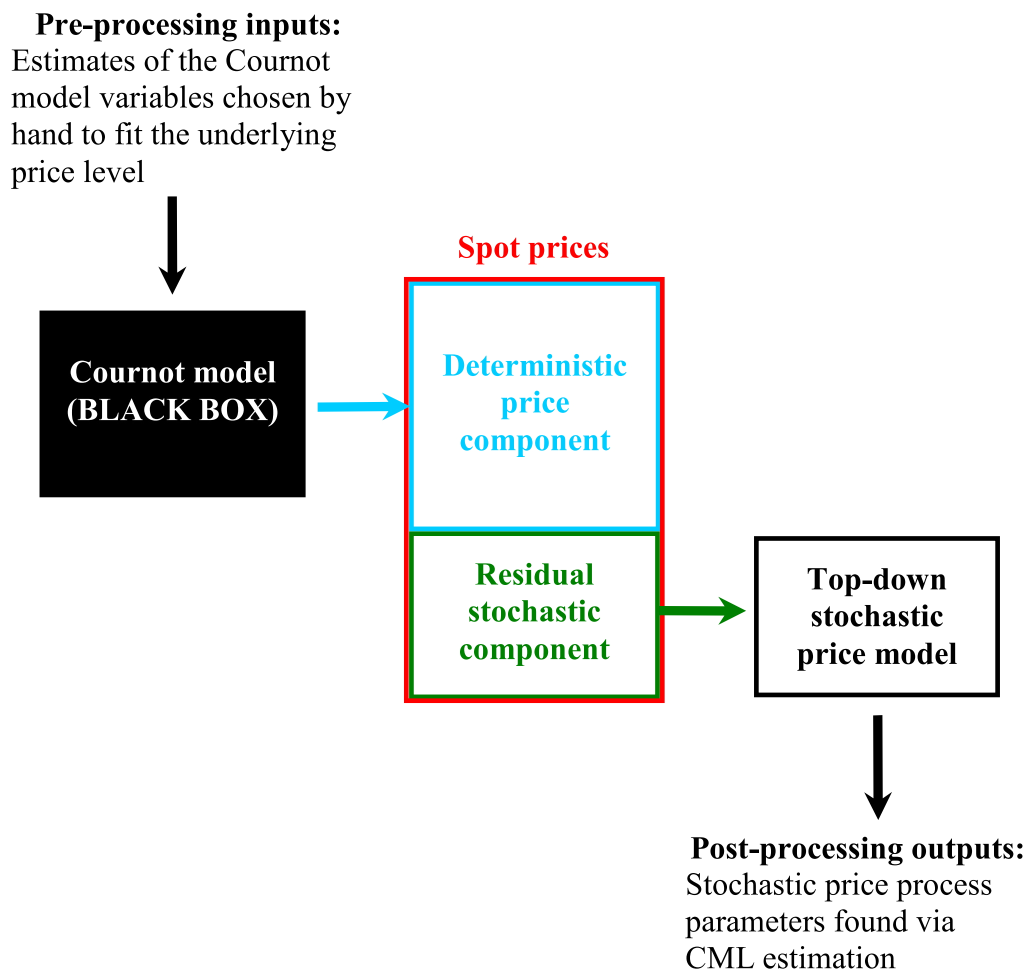

Figure 8.7: Diagrammatical representation of the pre- and post-processing methodology

8.5 Results

8.5.1 Fitting the EPV model to the peak, shoulder and off-peak series

The estimated parameters for the EPV model for each of the three Victorian price series can be seen in Appendix I. The composition of the deterministic component is different for each of the three series. As expected, the peak series has the highest constant (i.e. January) term, followed by that of the shoulder series. While the peak and off-peak series both have significant and increasing long term trends, the trend for the shoulder series is not statistically significant. Also, the peak trend is three times greater than the off-peak trend. The increasing trends may suggest an overall increase in load or generation costs, which would be felt more heavily when the more expensive and less efficient generation is required to be dispatched at peak times, hence the reason the trend in peak prices is greater. Interestingly, each of the monthly dummy variables is statistically significant in the shoulder series, only one is significant in the off-peak series, and about half of those in the peak series are significant. This suggests that there is some variation between the underlying level of prices throughout the year in the peak series, a great deal of variation in the shoulder series, but almost no variation in the underlying level of the off-peak series.

The estimated jump distributions of the three series are very different. The peak series has jumps on approximately 17 days a year, and these are normally distributed with a mean size of \$330 and a standard deviation of \$640. The shoulder and off-peak series have many more jumps, approximately 38 per year; however these have a much smaller expected magnitude and variance. This result is as expected, as the overall system is much less stretched in the lower load periods, and jumps in these periods are more likely due only to supply-side or transmission factors which will not have as serious an effect on prices as the major load-induced jumps in the peak series. That is not to suggest that factors other than load do not result in price spikes in the peak series; however volatility in the shoulder and off-peak series is less likely to be induced by extreme fluctuations in load. This would also explain the much greater unconditional variance (represented by ω) experienced in the peak series than in the other two series.

Possibly because the majority of peak volatility is load-induced, and load tends to fluctuate quite markedly from day to day in the peak periods of the NEM, the estimated autoregression parameter for the stochastic component in the peak series has a lower magnitude than those of both the shoulder and off-peak series. This means that shocks to the price series do not last as long in the peak series, and that prices revert to their deterministic level more quickly. This is understandable, particularly when the case of price spikes is examined. Price spikes in the NEM are often the result of a combination of extreme events occurring simultaneously, such as high load and supply outages or transmission constraints becoming binding. The probability of these events co-occurring is relatively small, and the probability of them occurring together for more than one day is almost zero. Therefore price spikes in the NEM are large but short-lived, compared to price rises in New Zealand which are less extreme but last a longer time.

The final observation that can be made regarding the parameters of the three stochastic components is that the composition of the GARCH conditional variance process differs between the off-peak and the other two series. In the off-peak series, there is a high degree of longer-term persistence in the volatility, represented by a statistically significant β (GARCH) term, which is not present at all in the other two series. In contrast, in the peak series the α (ARCH) term is statistically significant, suggesting significant short-term price volatility, while in the shoulder series neither α nor β are statistically significant. This further reinforces the fact that volatility is short-lived in the peak series, but more persistent in the off-peak series.

8.5.2 Cournot model calibration

The calibration results are separated into three sections, one for each of the three series that were analysed, while the peak section contains the results for both the weekday and weekend peak periods. Each section contains a table of the mean MADs that resulted from each simulation, with each simulation having as input a different combination of PED and PCL. As the values for the PED used were not particularly close together, simulations were run for each of the seven PCLs to narrow down the range of elasticities that produced the minimum mean MADs. The search was ultimately narrowed down to find the PED to two decimal places that yielded the lowest mean MAD for each PCL. Each section contains analysis of the results of each series, while a discussion of the three series' results follows in Section 8.6.

8.5.2.1 Peak

Table 8.1 contains the weekday peak results, which show that for each PCL there is a range of PEDs in which the MAD is minimized. For the higher PCLs (75%-90%) the

MAD is minimized when the PED is more inelastic, and for the lower PCLs the MAD is minimized when PED is more elastic. This is consistent with two explanations of the (relatively) low level of peak prices – either demand is inelastic, which would be in line with conventional wisdom regarding peak periods, and contract levels are high[11], or demand is elastic and contract levels are low. Figure 8.8 plots the change in MAD for each PCL while varying the PED, and shows that for the particular PCLs tested, the lowest MAD occurs with a PCL of 90% and a PED of -0.06. This leads us to believe that the first explanation may be true: demand *is* relatively inelastic in peak periods, which, as discussed earlier, could lead to higher prices and/or more abuse of market power. But, generators are, or at least act as though they are, highly contracted at such times – possibly around 90%, but definitely between 100% and 75%. The PCL is definitely below 100% though, as, regardless of the PED, PCLs of 100% and 110% underestimate prices overall. Prices are therefore above their perfectly competitive level for the majority of the time.

	PCL						
PED	**110%**	**100%**	**90%**	**75%**	**50%**	**25%**	**0%**
-2	10.21	10.26	10.22	10.20	10.25	10.42	10.45
-1.5	10.11	10.20	10.15	10.14	10.29	10.35	10.46
-1	9.85	9.89	9.88	10.00	9.90	10.15	10.16
-0.5	9.90	9.85	9.84	9.73	9.68	10.19	11.53
-0.3	10.42	10.34	9.98	9.50	9.94	12.79	16.48
-0.15	13.35	11.65	9.85	9.40	15.48	24.80	35.61
-0.05	22.72	14.07	9.11	22.47	53.35	90.29	126.58
-0.01	36.21	15.87	47.14	148.03	325.98	511.29	709.29

Table 8.1: Mean Absolute Deviations (in $AUS/MWh) for the simulations of the weekday peak series prices

[11] As mentioned earlier, it is also possible that the contract levels are implicit indicators of the threat of regulation. When demand is inelastic, generators are aware that their behaviour may be under scrutiny and, as a result, they may act as though they are heavily contracted.

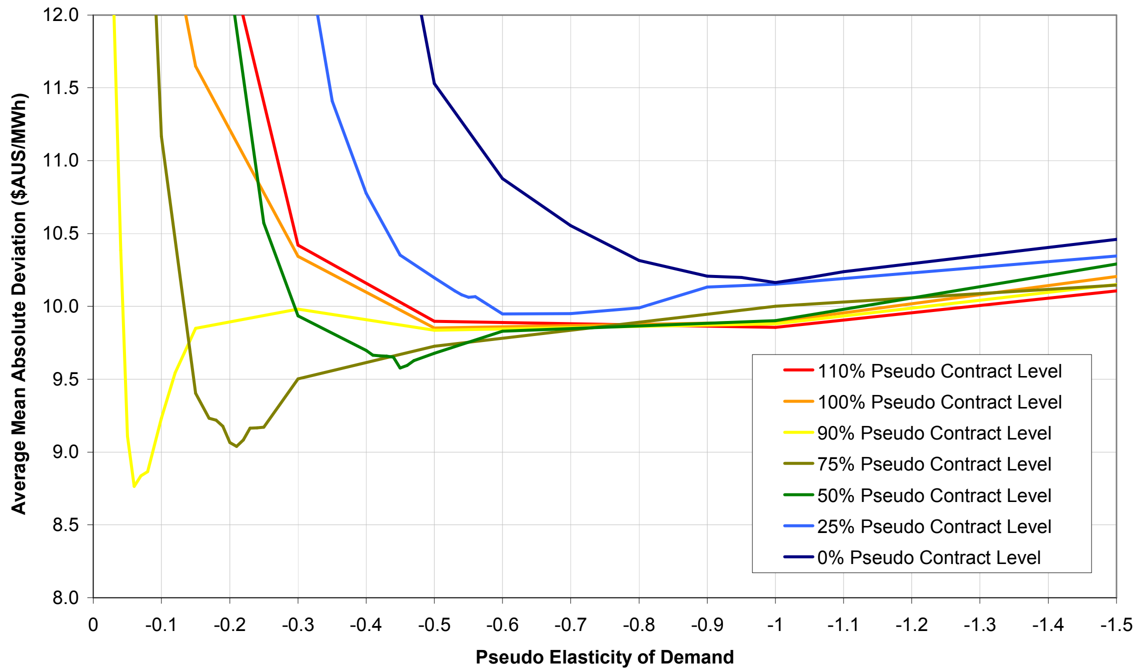

Figure 8.8: MAD versus PED while holding the PCL constant: weekday peak series

The results for the weekend (and public holiday) peak period are shown in Table 8.2, and tell a completely different story. The minimum MAD is achieved with a PCL of 0% and a PED of -0.52, and the exact PCL that minimises the MAD would be, at the most, 25%. This suggests that demand is particularly elastic in the weekend peak periods, with consumers appearing happy to consume less power if prices increase. The high elasticity may also suggest that as peak weekend loads are lower than peak weekday loads, there is likely to be a greater amount of excess capacity, and therefore generators' behaviour will be more aggressive. For the lower levels of load, the fight for a share of the market will be more competitive as each generator faces a more elastic residual demand. The low PCL goes hand-in-hand with the high PED, for the reasons mentioned earlier: there is less opportunity for generators to exercise market power when elasticity is greater, there is a lower threat of regulation, and they therefore have less reason to act as though they are contracted.

	PCL						
PED	**110%**	**100%**	**90%**	**75%**	**50%**	**25%**	**0%**
-2	8.49	8.49	8.49	8.46	8.28	8.10	7.89
-1.5	8.98	8.98	8.93	8.88	8.75	8.42	7.87
-1	10.00	9.93	9.79	9.60	9.10	8.19	7.59
-0.5	11.87	11.68	11.23	10.30	8.31	6.79	5.97
-0.3	13.60	12.95	12.00	9.76	6.81	6.54	8.26
-0.15	16.09	14.26	11.36	7.17	7.90	12.95	18.64
-0.05	20.74	15.01	7.00	12.10	30.20	49.31	70.29
-0.01	30.90	15.12	28.05	81.70	187.83	293.26	402.54

Table 8.2: Mean Absolute Deviations (in $AUS/MWh) for the simulations of the weekend peak series prices

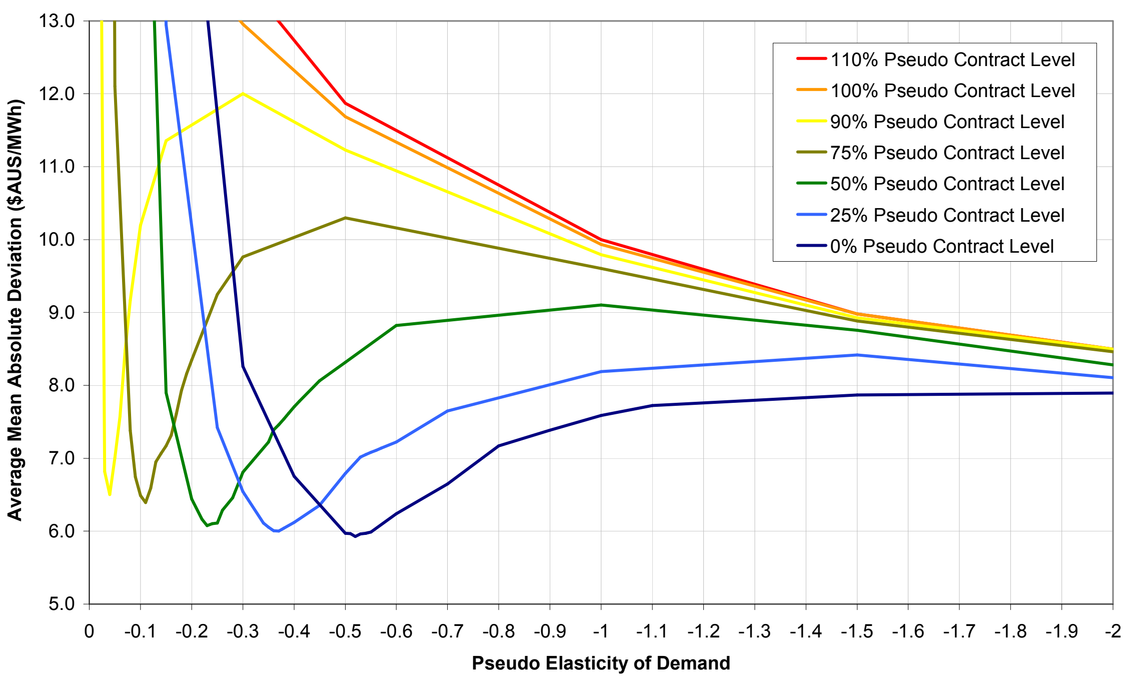

Figure 8.9: MAD versus PED while holding the PCL constant: weekend peak series

From Figure 8.8 and Figure 8.9 it can be inferred that for each PCL there is a range of PEDs that results in simulated prices being too low, a PED for which the MAD is minimized, and a range that leads to simulated prices being too high. For each curve on these charts corresponding to a specific PCL, the part of the curve to the left of the minimum point corresponds to simulated prices generally being greater than the actual price level. This leads to large negative residuals (i.e. large negative errors or large absolute deviations). Conversely, to the right of the minimum point prices are in general underestimated, leading to large positive residuals and absolute deviations. However, the amount that prices can potentially be overestimated (i.e. through a highly inelastic PED and low PCL) is much lower than the amount prices can be underestimated (through an elastic PED and high PCL).

This is shown in Figure 8.10, which plots two different simulated peak series, both with a PCL of 90%, but having different PEDs. The series with more inelastic demand (PED = -0.01) consistently overestimates the underlying price level, while the other series (PED = -0.15) underestimates. Another interesting feature of the two graphs is that the range of PEDs in which simulated prices are a relatively good fit to actual prices is smaller for the higher PCLs (75% to 90%) than for the lower PCLs (0% to50%). This makes it much easier to pinpoint the PED that minimizes the MAD for higher PCLs, which is seen with the shoulder and off-peak periods' results as well. However, the results of the simulations with lower PCLs are much more sensitive to changes in the PED at inelastic levels, with the combination of highly inelastic PED and low PCLs leading to enormous simulated prices.

Also, note that the MADs for each of the different PCLs converge on a single point as the PED becomes more elastic. The PCL becomes irrelevant as the PED is more elastic, which makes intuitive sense as there is less need for contracts if there is virtually no potential for the abuse of market power. This observation is explored further in Appendix H.

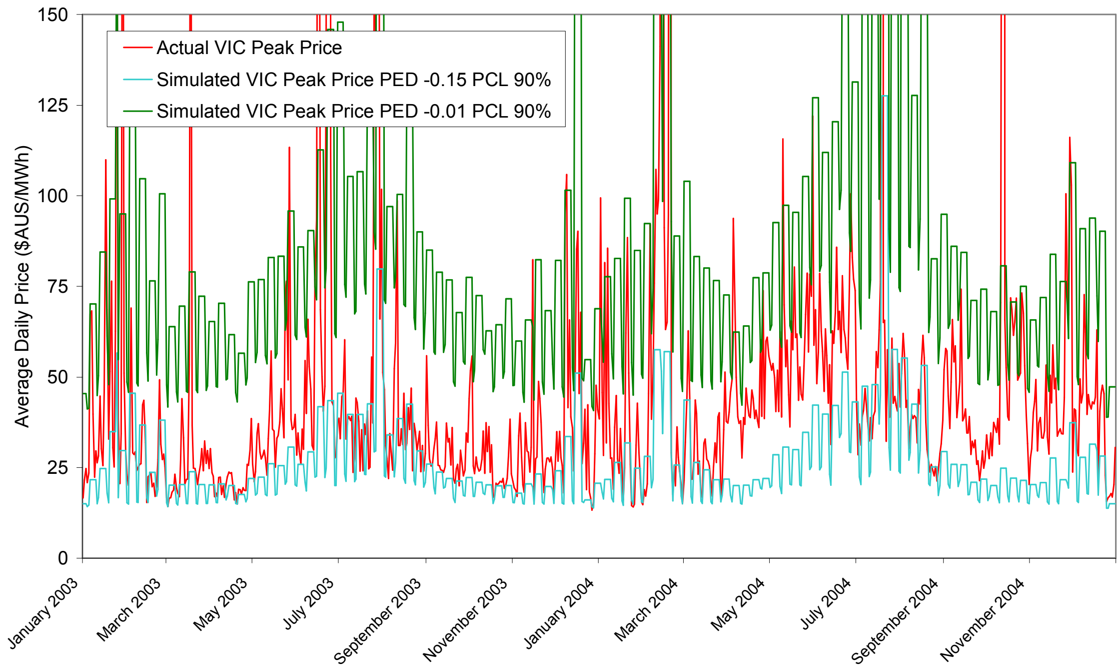

Figure 8.10: Average daily spot price in Victoria for each day of the peak series, 1 January 2003 – 31 December 2004, and two simulated price series resulting from a PCL of 90% but different PEDs

The series with the lowest MAD, generated with a PCL of 90% and a PED of -0.06 for the weekdays and 0% and -0.52 for the weekends, is plotted in Figure 8.11 alongside the actual peak price series. Initial examination of the two series shows that the fit of the underlying price level is fairly accurate in 2003 but much less so in 2004. Apart from July and August 2004, prices are consistently underestimated by the simulation throughout the year. Without a more thorough knowledge of the actual market events in 2004 it is impossible to know why this underestimation occurs when the fit is so good in 2003. However there are several possible explanations, each of which would render using a constant set of input data to represent both years inappropriate. As the estimated EPV model parameters included a statistically significant trend, it is possible that this upward increase in prices was the result not of increasing load over time, which would have been accounted for by T-CONE, but possibly some other factor such as increasing fuel costs or more lenient regulation. Such a discrepancy between the simulated results and actual prices was always likely with so many variables (including fuel costs, contract levels, capacity, etc.) held constant over the course of this simulation. As is explained later in this chapter, it turns out that the output levels from several of the NEM's major base-load generators were fairly variable throughout 2004, which could not be accounted for in T-CONE's simulations .

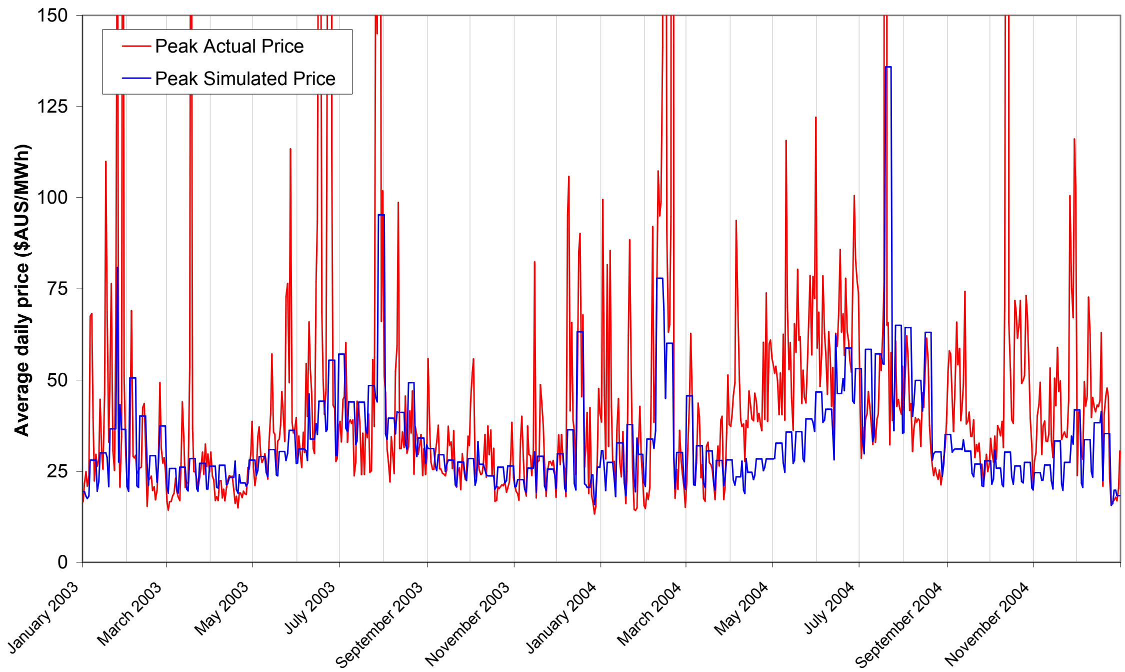

Figure 8.11: Average daily spot price in Victoria for each day of the peak series, 1 January 2003 – 31 December 2004, and the simulated price series resulting from a PED of -0.06 and a PCL of 90% for the weekdays, and -0.52 and 0% respectively for the weekends.

8.5.2.2 Shoulder

The shoulder simulation results, shown in Table 8.3 and Figure 8.12, tell a similar story to those from the weekday peak series. Again, the MAD line for each PCL below 100% appears bimodal. The range of PEDs over which the MAD is approximately minimized is larger with the lower PCLs. And again, as the PED becomes more elastic, the PCL has less influence.

The combinations of PCL and PED that minimize the MAD are similar to the weekday peak series results, except that for each PCL the minimum MAD is found with a more elastic PED than before. Conventional wisdom tells us that demand should be more elastic than in the peak periods, however one thing that is certain is that generators' behaviour will be more competitive in the periods with lower loads. As shown on Figure 8.12, the combination that minimizes the MAD is a PCL of 75% and a PED of -0.25, although the minimum MAD for a 90% PCL is fairly similar.

	PCL						
PED	**110%**	**100%**	**90%**	**75%**	**50%**	**25%**	**0%**
-2	4.86	4.86	4.85	4.83	4.78	4.75	4.69
-1.5	4.82	4.80	4.79	4.79	4.71	4.70	4.66
-1	4.80	4.77	4.70	4.61	4.49	4.48	4.55
-0.5	5.21	4.99	4.74	4.23	4.23	4.91	6.00
-0.3	6.46	5.67	4.75	3.94	4.98	6.97	10.00
-0.15	8.73	6.82	4.40	4.63	9.17	16.08	23.35
-0.05	14.20	7.55	5.07	15.09	36.15	59.10	81.50
-0.01	24.12	7.80	33.86	96.13	210.89	324.60	443.45

Table 8.3: Mean Absolute Deviations (in $AUS/MWh) for the simulations of the shoulder series prices

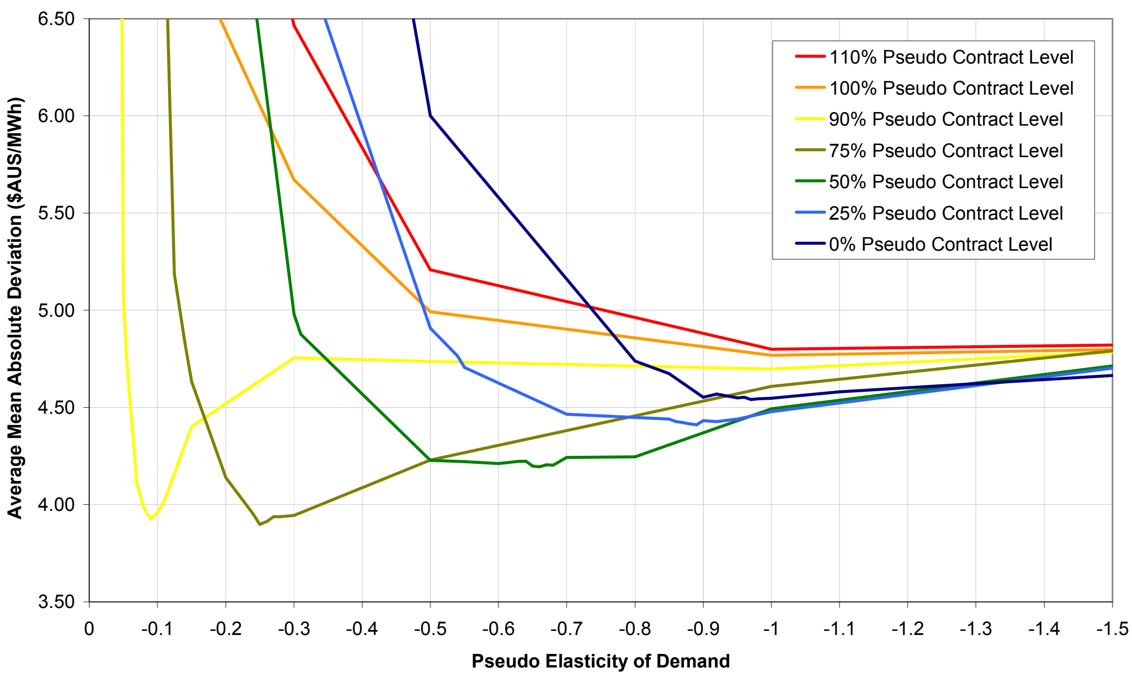

Figure 8.12: MAD versus PED while holding the PCL constant: shoulder series

The fit of the simulated series with a PCL of 75% and a PED of -0.25 is shown in Figure 8.13. Again, the fit appears good for most of 2003, but not as good for most of 2004 apart again from July and August. In those two months the load was obviously higher, as can be seen in Figure 8.4, and the simulated prices rise to reflect that. However, there was no load-driven reason why prices should have risen to that level earlier, at around the start of May. Like the peak series, the prices are underestimated much of the time in 2004.

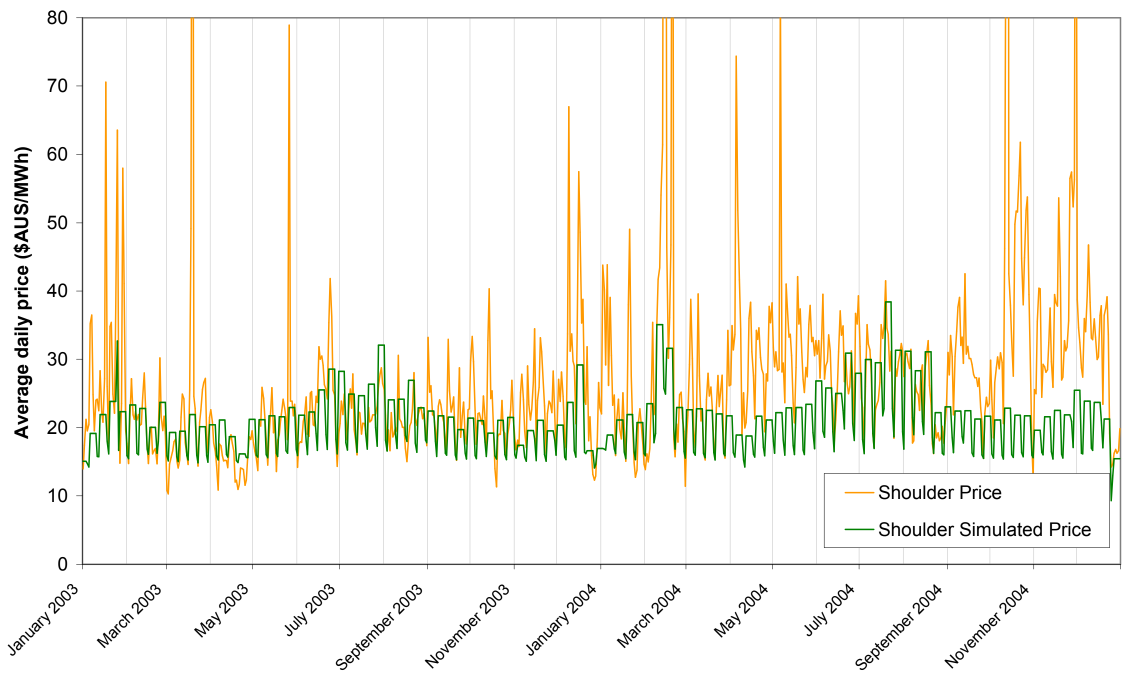

Figure 8.13: Average daily spot price in Victoria for each day of the shoulder series, 1 January 2003 – 31 December 2004, and the simulated price series resulting from a PED of -0.25 and a PCL of 75%

8.5.2.3 Off-peak

The results for the off-peak series, shown in Table 8.4 and Figure 8.14, tell a different story to the other series' results. Apart from the 90% PCL simulations, the fits of all the other PCLs improve as the PED becomes more elastic.

	PCL						
PED	**110%**	**100%**	**90%**	**75%**	**50%**	**25%**	**0%**
-2	2.62	2.62	2.61	2.54	2.51	2.51	2.53
-1.5	2.64	2.64	2.61	2.55	2.52	2.54	2.63
-1	2.75	2.72	2.69	2.56	2.55	2.67	2.86
-0.5	3.10	2.93	2.75	2.56	2.78	3.31	4.19
-0.3	3.40	3.11	2.67	2.64	3.47	5.11	7.28
-0.15	4.01	3.20	2.58	3.39	7.05	11.75	16.62
-0.05	6.68	3.24	3.76	11.51	25.83	39.75	54.34
-0.01	16.37	3.27	25.21	62.11	131.45	202.95	275.20

Table 8.4: Mean Absolute Deviations (in $AUS/MWh) for the simulations of the off-peak series prices

The best fits for each PCL occur when the PED becomes more elastic, but the increase in fit is very slight. In fact, the minimum MAD for each PCL is virtually the same, except for PCLs 110%, 100% and 90%. It is logical to assume that generating companies will not be fully or over-contracted in off-peak periods. We assume the reason that the best-fitting results for each PCL were so similar was that as the load in the off-peak period is relatively low compared to total capacity, opportunities to exercise market power are almost non-existent. Generators' bidding behaviour would be very competitive in order to capture a share of the low load, creating a high elasticity in the residual demand curve faced by each competitor. Also, the demand curve would intersect the aggregate supply curve at a low and very flat point of the supply curve each day, most of the load will be met by base-load generators with similar marginal costs, and, as a result, the simulated price will not vary very much at all as contract levels and elasticities change.

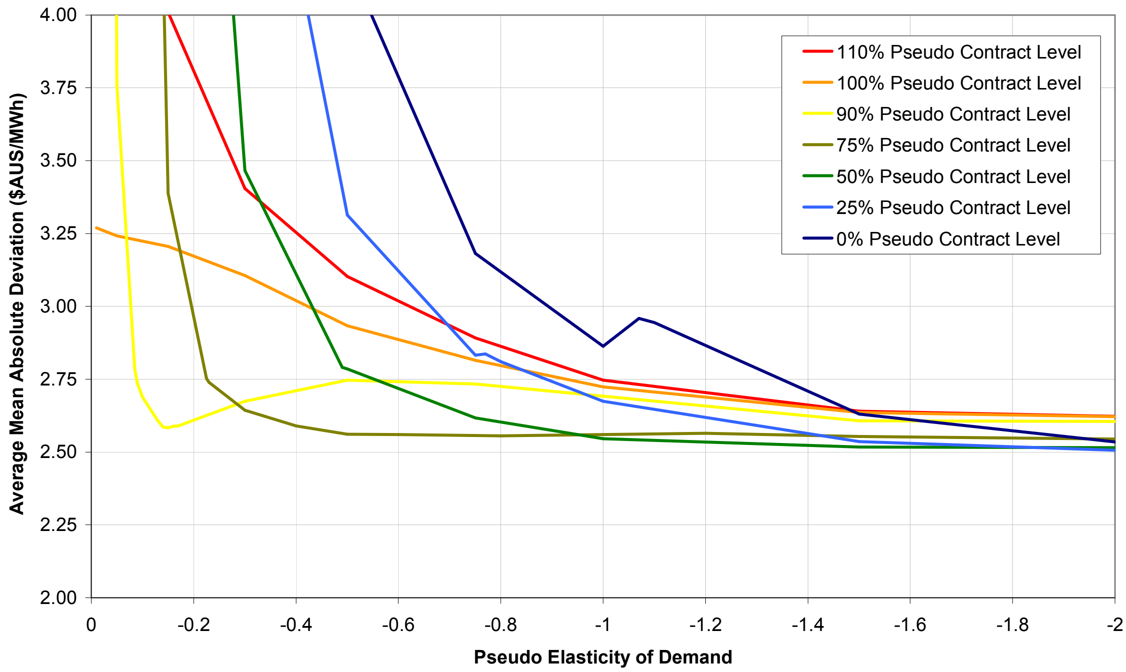

Figure 8.14: MAD versus PED while holding the PCL constant: off-peak series

As a result of the similarity between these results we did not narrow down our search for the minimum MAD between the points on the initial grid as thoroughly as with the other two series, and decided to choose the series with the minimum MAD on the grid as the best fit. This was the combination of a PED of 2 and a PCL of 25%, and the entire simulated price series is shown in Figure 8.15.

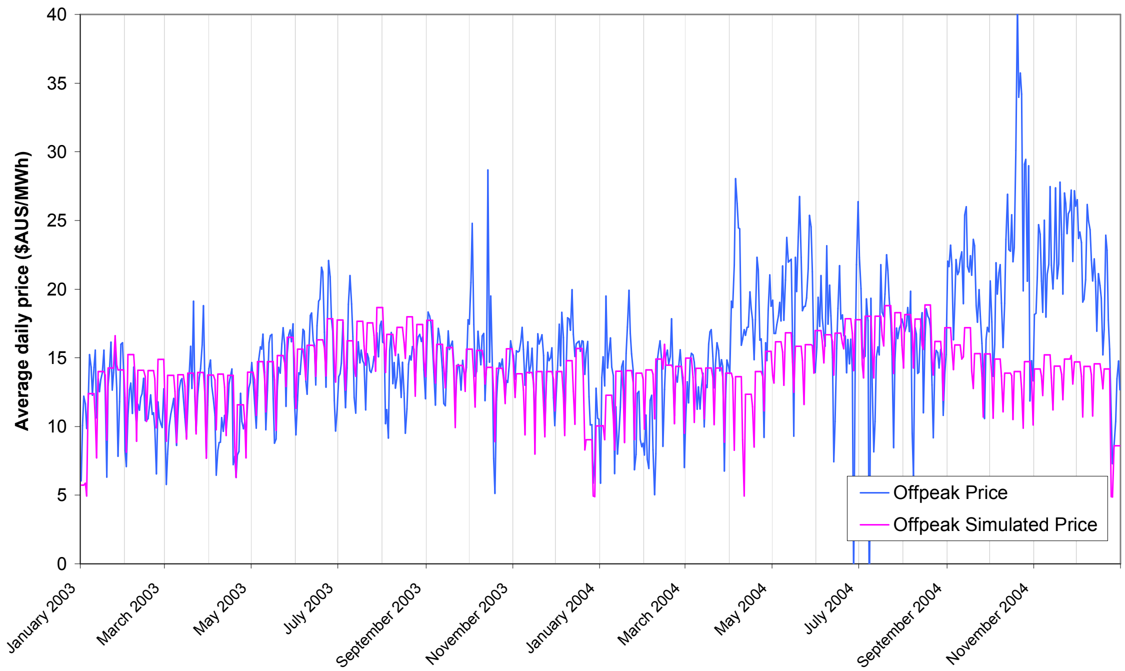

Figure 8.15: Average daily spot price in Victoria for each day of the off-peak series, 1 January 2003 – 31 December 2004, and the simulated price series resulting from a PED of -2 and a PCL of 25%.

Once again, the fit is relatively good for 2003, but poor for most of 2004. The degree of underestimation of the 2004 prices is more obvious with this series. As shown in Figure 8.4, total system load was much lower in spring 2004 than it was in winter, which is reflected by the drop in the simulated prices, therefore the increase in prices was definitely not load-driven. Also, the prices in all four regions increased from September onwards, ruling out a shortage of inter-regional transmission capacity, so the price rise must have resulted from a supply-side factor.

As load is effectively the only variable input into the T-CONE model, it is worth plotting a graph of daily load versus the simulated Cournot price for each load period, using the final values for the PED and PCL selected. This graph is shown below in Figure 8.16. The off-peak prices are virtually monotonically increasing in load, which is as expected due to the fact that transmission constraints are unlikely to be binding anywhere on the network at these times. The relationship between load and price in the other two load periods is much less clearly defined, due to the interaction of higher loads across the network and also the greater occurrence of binding transmission constraints.

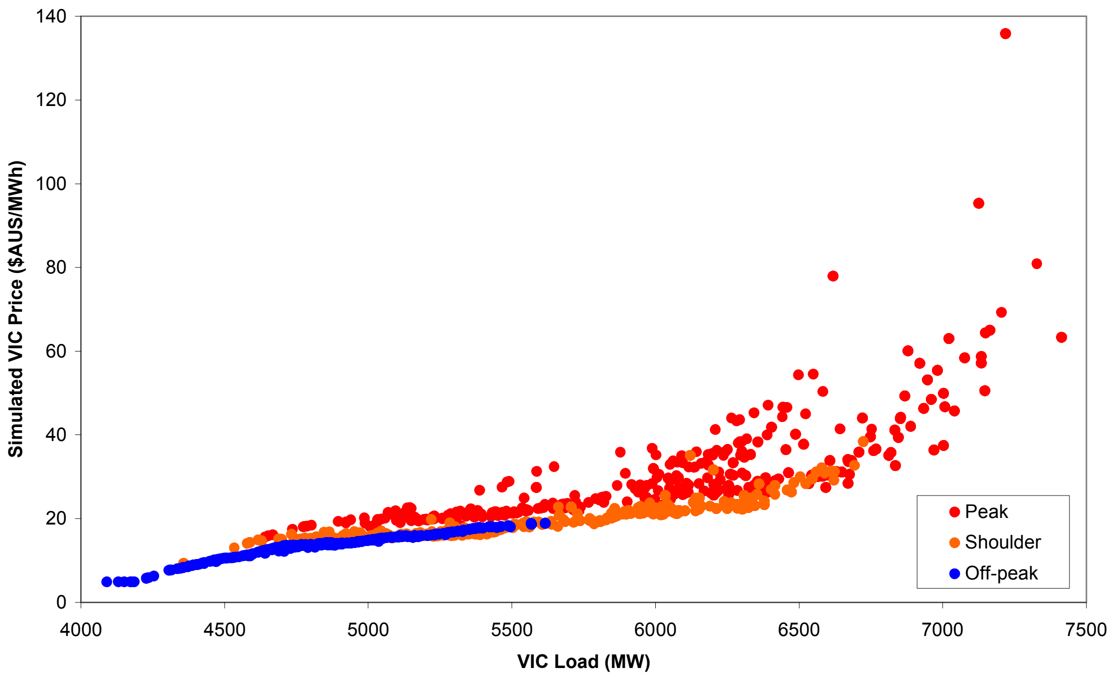

Figure 8.16: Load versus simulated price in Victoria for each day of the peak, shoulder and off-peak series, 1 January 2003 – 31 December 2004, using the final values of the PED and PCL selected for each of the three load periods

While it is obvious that incorporating load into a top-down model for the NEM prices may have produced similar results, such a model is unable to forecast generator dispatch and network flows, and thus serves a much more limited purpose than the Cournot model.

8.5.2.4 Estimated parameter interpretation

The range of PEDs for each of the different load periods lines up with conventional wisdom. During peak periods, prices are highly sensitive to shifts in demand, and hence a large shift in the price may not yield a very great shift in demand. It could be that demand really is inelastic in peak periods, but it is definitely the case that the part of the supply curve in which the marginal generator is found in peak periods is much steeper compared with in off-peak periods. Also, recall that the PED in a Cournot model serves as a measure for generator response. In off-peak periods, when loads are lower, generators have much more excess capacity than in peak periods. They therefore each have to bid much more competitively to ensure they are dispatched in off-peak periods, and naturally face a much more elastic residual demand curve than in peak periods.

The estimated contract levels are less intuitive. Generator behaviour in peak periods appears to be highly constrained by (implicit) contracts. This means that they do not push prices up as high as they might, possibly because of the threat of regulatory action. However, while contracts in the NEM can vary by peak / off-peak, there is no straightforward reason why generators should act or feel as though they are less contracted in the lower load periods. They may be less concerned about contracts because elasticity is high and therefore price volatility is low, but the amount of offered capacity varies more widely in off-peak periods than it does in peak periods. The main priority for generators is to have their plants running during the peak periods when prices are high. Therefore, they are more likely to take their plants down in off-peak periods if they need short-term maintenance, which could explain behaviour representing a lower level of contracts in off-peak periods. Alternatively, the NEM contains a number of large must-run thermal units which, in order to ensure they are running in peak periods, must be kept running in off-peak periods and bid low, often negative prices to ensure they are dispatched. This behaviour would be interpreted as though the company were highly contracted in the off-peak period, when in fact they were not contracted at all.

8.5.3 Fitting the stochastic process model to the residual price series

After the simulated series yielding the lowest MAD has been found for each period, the three residual series can be calculated. For each of the three series, this involves subtracting the simulated price for each day from the actual price for that day, resulting in a series of 731 residuals for each of the peak, shoulder and off-peak series. The parameters of the stochastic component of the EPV model are then fitted to each of these three residual series in turn. The estimated parameters are listed in Appendix I, and a discussion of each series' results appears in its own section below.

The peak, shoulder and off-peak residual series are plotted in Figure 8.17, Figure 8.18 and Figure 8.19 respectively. Each figure shows residuals clustered around zero (aside from the price spikes) for much of 2003 and the first few months in 2004, but, as discussed above, each residual series is well above zero for most of April-May and September-December 2004.

8.5.3.1 Peak

The peak residual series is shown in Figure 8.17, with the residual (y) axis truncated. The major comment on the estimation results is that the fit of the stochastic component to the residuals is worse (as measured by the log-likelihood) than the combined fit of the full EPV model in section 8.5.1. This suggests that there is no improvement to the fit of prices when a Cournot model is used instead of the deterministic component of the EPV model. The jump distribution parameters have not changed markedly, indicating the Cournot model does not do any better or worse at modelling the price spikes. The β term has become significant, indicating a degree of longer-term persistence in the volatility that was not present in the whole EPV model.

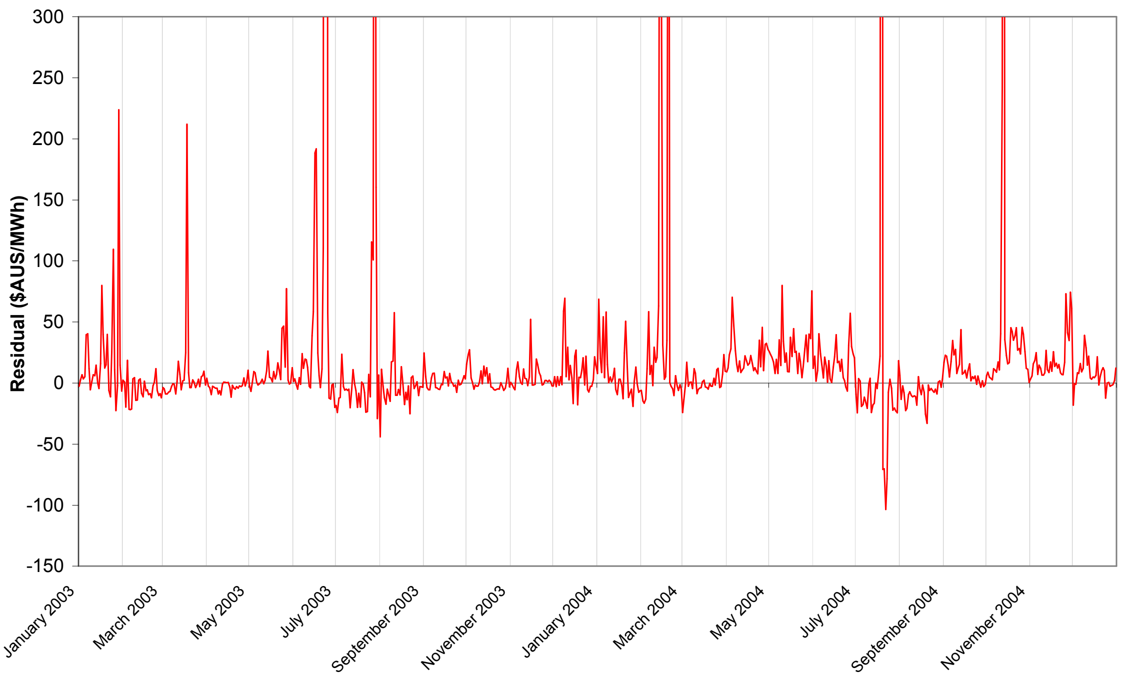

Figure 8.17: Series of peak residuals not modelled by the Cournot model, 1 January 2003 – 31 December 2004

In order to determine why the introduction of the Cournot model decreased the fit of the price model, the fit of the whole EPV model (deterministic component included) to the peak residual series is assessed as well. The estimated parameters are in Appendix I also. Five of the deterministic component parameters are statistically significant – the trend, May, July and August monthly dummy variables and the weekday dummy variable. None of these three monthly dummy variables were significant in the original estimation in Section 8.5.1, and none of the four monthly dummy variables that were significant in the original estimation were significant in this one. It could be expected that the trend is significant, due to the fact that the residuals in the second half of the sample period are in general greater than zero, but the trend value is only 61% of the size of the initial trend value. This suggests that around 40% of the positive trend was due to increasing load (which T-CONE could account for), with the remaining 61% due to some other reason.

The important statistics to note from the fit of the overall EPV model to the peak residuals are the Log-Likelihood and Schwartz Information Criterion[12], both of which indicate that including both the EPV deterministic component and the T-CONE prices into the overall price model gives a better fit than just using the EPV model. It is reassuring to know that T-CONE does improve the price modelling process overall.

8.5.3.2 Shoulder

In contrast to the fit to the peak residuals, the fit of the EPV stochastic component to the shoulder residual series is an improvement over the fit of the whole EPV model to the shoulder price series. This is despite the shoulder residuals being greater than zero for much of 2004, as shown in Figure 8.18. The estimated jump distribution is markedly different to that reported in Section 8.5.1, with the estimated jump intensity more than halved, and the mean jump size and variance greatly increased. The unconditional variance has also more than doubled, and, interestingly, the autoregression parameter has increased. This suggests a shift from many small, short-lasting jumps as estimated in Section 8.5.1 to a smaller number of large jumps (and more overall volatility) with a longer-lasting effect. If it is assumed that the stochastic component accounts for movement away from and back towards zero in the residual series, this assessment of the stochastic component parameters looks in Figure 8.18 to be accurate.

[12] Recall from Chapter 3 that these statistics are measures of the goodness-of-fit, based on the likelihood of the estimated model parameters being correct given the prices actually observed.

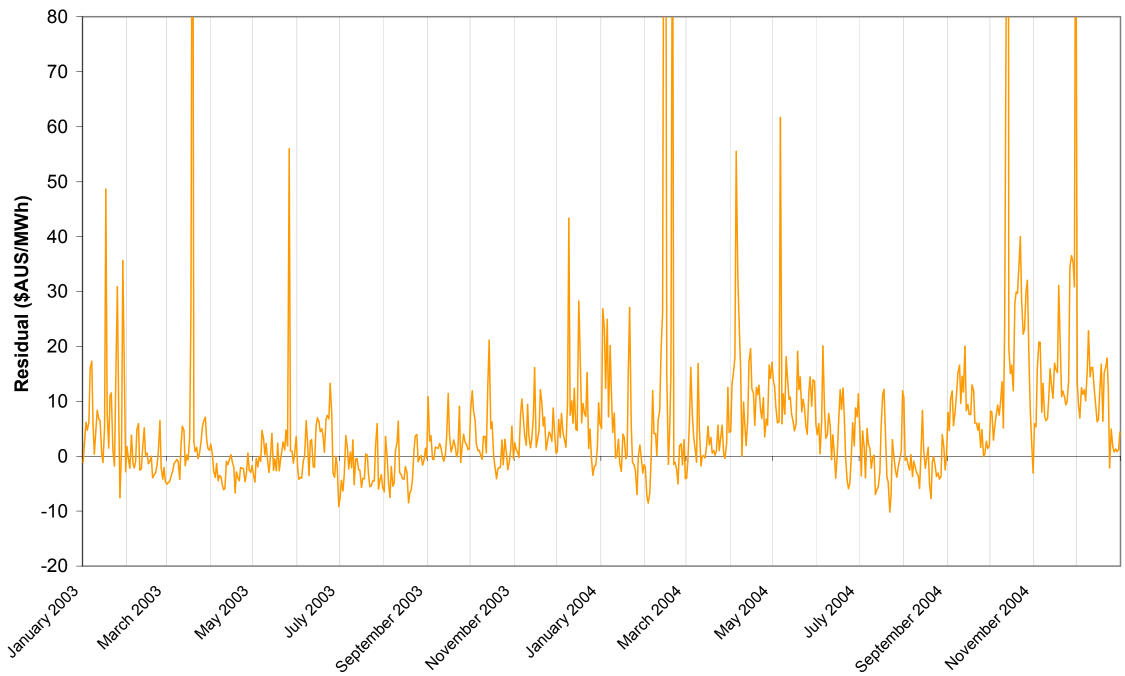

Figure 8.18: Series of shoulder residuals not modelled by the Cournot model, 1 January 2003 – 31 December 2004

8.5.3.3 Off-peak

As with the shoulder series, the fit of the EPV stochastic component to the off-peak residual series (which is shown in Figure 8.19) is also an overall improvement on the fit of the entire EPV model to the off-peak price series. Also, like the estimated parameters for the shoulder residuals, the estimated nature of the jump distribution is different, with the estimated jump probability being lower and the estimated jump size variance larger. Otherwise, there are no other notable changes in the nature of the estimated parameters.

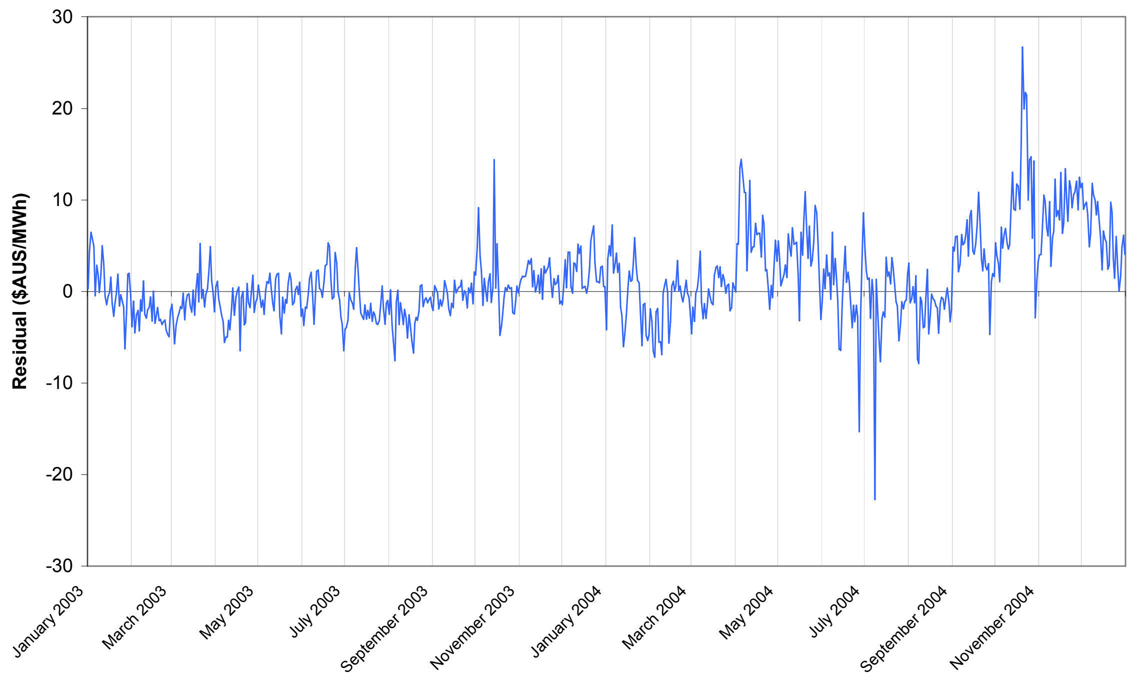

Figure 8.19: Series of off-peak residuals not modelled by the Cournot model, 1 January 2003 – 31 December 2004

Bushnell (personal communication, 2007) writes that he thinks "one of reasons why the Cournot model is doing a better job off-peak is the representation of the contracts. My understanding is that "cap" or one-way CFDs are very common instruments in Australia. Such a contract has the effect of impacting the Cournot equilibrium only around the strike price (see Willems, 2006). Off peak, when equilibrium prices are well below the strike price anyway, these contracts will have little effect. However if they are in play during the peak periods, they will have a binding impact. So if firms possess a bundle of two-way CFDs *combined* with one-way CFDs, we would see the pattern described in the data. This fits the stylized institutional facts about this market".

However, from our experience there are not actually enough one-way CFDs in the NEM to make such a difference to the Cournot equilibria in the peak period. It is much more likely that the slightly inferior fit of the combined peak price model was due to other factors, such as generator outages, as discussed in the next section.

8.5.4 Why did the Cournot model underestimate prices for 2004?

For each of the peak, shoulder and off-peak periods, the simulated prices using the best-fitting combination of the PED and PCL underestimate the actual price level for much of 2004. One of the possible reasons for this conjectured earlier in this section is that the amount of generating capacity available in 2004 was not constant throughout the year. An examination of the actual generation data for each plant (again sourced from CRA International's NEM database) confirms that this is the case.

The simulated prices in October and November in 2004 are much lower than the observed prices. Interestingly, the average load observed during these months was lower than in both September and December, but spot prices were higher and more volatile. This confirms that the increase in the price level was not load-driven, and therefore could not be accounted for in forecasts by T-CONE keeping the supply mix constant.

With regards to the generation levels of the individual power stations with the cheapest SRMC (estimated at below AUS$10/MWh), both Loy Yang A and Loy Yang B were generating at around 500MW below their respective observed capacities for much of that period, but at alternating times. Yallourn decreased generation by as much as 700 MW during that time, and Millmerran halved its production (a reduction of around 430MW) for almost all of October. Of the slightly more expensive generators (SRMC below $20/MWh), Bayswater reduced generation by around 500MW for virtually all of October and November, and Munmorrah did not operate at all for most of October and November (it had been operating at 300MW in September. This means that at times, up to 2500MW of base-load generation was not operating, for reasons not investigated in this thesis.

The prices produced by T-CONE also underestimate the actual price level for all of April and May 2004. After March, there was a large drop in load in April, and while loads in May were similar to March, actual prices were much higher in both April and May than they were in March. The main reductions in base-load generation after March were at

Yallourn (around 700MW), Millmerran (430MW), Bayswater (500MW) and Tarong North (around 400MW).

During both autumn and spring in 2004, over 2 GW of base-load generation did not operate, however in T-CONE this was not accounted for, with generating capacity remaining constant throughout. The simulated prices for the off-peak period, shown in Figure 8.15, illustrate this well. Load was lower in these seasons than in winter, as shown by the lower prices from T-CONE, however the actual spot prices were much higher and more volatile. Even without looking further into this supply-side discrepancy, it can be concluded without doubt that T-CONE would have been able to model prices more accurately in those periods had the input data been adjusted to take these capacity changes into account.

8.6 Conclusions

8.6.1 Modelling spot prices

The overall conclusion that can be drawn from this exploratory study is that using a Cournot model to model the deterministic level of electricity spot prices can improve the overall modelling fit of a complex time series price model. Now calibrated with real market data, the Cournot model also has the advantage over other price forecasting models of being able to calculate generation levels and transmission flows.

8.6.2 Estimates of the PED and PCL

In terms of the Cournot calibration, the ranges in which the estimated PEDs lie for each period seem plausible according to conventional wisdom. The estimated PCLs also line up with conventional wisdom, with a higher level in the peak periods and a lower level in the shoulder and off-peak periods. However, it may be that instead of conventional wisdom actually being correct, this wisdom has been thought to be correct for long enough for it to determine behaviour.

The fact that the realised PCL for the off-peak period could not be estimated conclusively is interesting. Off-peak prices are in general low, with low volatility; therefore, because of the low level of risk, consumers do not have a major incentive to buy contracts. However, the NEM contains a number of major thermal units, which often bid very low (sometimes negative) prices in a struggle to keep generation up to minimum levels overnight, thus operating at a loss so as to be available to generate profitably the next morning. Arguably, these might be modelled as behaving as if they are contracted to generate at a high percentage of their perfectly competitive levels, even though they may not be contracted at all. This would lead to some confusion in the apparent contract rate observed in off-peak periods.

It is interesting that the Cournot model on its own improved the deterministic fit of prices for the shoulder and off-peak periods, but not for the peak period. We suggest that due to the greater amount of excess capacity available in the non-peak periods, the prices are a great deal less sensitive to any changes in the supply mix that may occur. The demand curves in these periods intersect the aggregate offer stack at the lower end of the stack where it is relatively flat and does not change a great deal throughout the year, hence calculating the price at those intersections is relatively straightforward. In the peak periods, the demand curve intersects the aggregate offer stack at the upper end where it is much steeper, hence the spot price is very sensitive to any changes in offered capacities and prices that may occur. As a result, a Cournot model that holds supply-side information constant throughout the year will likely be much less accurate in its prediction of peak prices than it will in predicting shoulder and off-peak prices[13].

8.6.3 Future application of a calibrated Cournot model

As mentioned earlier, the combination of a calibrated Cournot model with a time series price model produces an improved fit to prices than the time series model can on its own.

[13] However, as noted in Section 8.5.4, reducing the amount of base-load generation still has an effect on shoulder and off-peak prices.

This therefore achieves one part of the goal set out at the start of this chapter, and gives a response to the claim made by Frame and Joskow (1998), who were "not aware of any significant empirical support for the Cournot model providing accurate predictions of prices in any market, let alone an electricity market." However, the Cournot model also has the advantage of being able to predict generation levels and flows across a network, which a time series price model cannot. Now that the Cournot model has been calibrated, and accurately reflects price behaviour in an existing market, its forecasting results for any application to a hypothetical market situation become much more credible.

Such situations may include the establishment of a completely new market, where the contract levels may be known but the elasticity of demand can be assumed from the results of this study, or where both the contract levels and elasticity of demand are unknown and must be assumed. Further, if a regulatory body wanted to know the potential effect on market prices in the NEM of splitting a large generating company or two generating companies merging, the calibrated Cournot model can give results more credible than a model which does not mimic the current market situation.

8.6.4 Directions for further research

While the calibration and inclusion of a Cournot model into a time series model for spot prices was a success in this exploratory study, there are obviously many changes to the methodology that would increase the accuracy of the process. As mentioned above, supply-side variables such as capacities and marginal costs do not remain constant over time, and would ideally be altered throughout the sample period. However, in a forecasting sense, it is hard to predict how these variables will change, so it is probably more appropriate to randomly vary factors like outage rates in a Monte Carlo simulation setting. Also, contract rates and elasticities vary across time and by region, so these could be also adjusted randomly or by season and/or region[14]. Further, some of the components

[14] There is an argument against varying the elasticity by season, however. We would expect the elasticity of demand in the winter shoulder period to be roughly equivalent to the elasticity of demand in (say) the autumn peak period. This is due to the fact that, in general, demand is greater in winter than in autumn

in the stochastic process could be linked to the simulated deterministic price level, in much the same way as the MWV determined the level of price volatility in the hydro price model presented in Chapter 5.

Overall, the success of this study opens the door for a much wider range of testing and calibration to be completed, and adds credibility to the use of these market models in practice.

---

and, correspondingly, the amount of excess supply will be less in winter than in autumn. However, it may in fact be the case that the generators optimise their maintenance schedule so that the amount of excess supply (and therefore the inherent threat of price-responsive competition) is constant throughout the year.

9

CONCLUDING REMARKS

The field of spot price modelling has become increasingly informed by new research, and the bottom-up and top-down models being developed can be complex. We have taken the point of view that each of these two types of models has positive and negative attributes, and in certain situations neither can do a satisfactory job of modelling the level and volatility of price series individually. This provided our motivation for combining the merits of both into a series of hybrid models.

The negative relationship between hydro storage levels and spot prices in New Zealand is well-known, but, to our knowledge, had not been quantified in a top-down price model. We used readily-available market data to both illustrate and quantify this relationship, using established theory on hydro reservoir management to estimate the value of water in the reservoirs. We showed that not only can the underlying level of spot prices be estimated using a concept we call the Relative Storage Level, but that the probability and likely size of a jump in prices can also be linked to the storage level. These are worthwhile additions to the academic literature, of which any future price model for the NZEM should take advantage.

In order to forecast NZEM price levels in the future, we required a model to forecast storage levels. Inflow modelling is a well established field of research; but no time series models of releases exist to our knowledge. We showed that the level of releases from New Zealand's reservoirs on any given day is also strongly linked to the value of the water in those reservoirs on that day, as well as being dependent on the inflows in the days preceding and following the day in question. This model of releases is able to forecast levels of release with a high level of accuracy.

When combined with the release model, the NZEM price model is able to forecast prices over a much wider range of inflow sequences than those observed in the limited lifetime of the market so far. This gives rise to many interesting applications of the combined model, particularly in predicting long-run behaviour of NZEM spot prices and analysing reservoir management policies in previous power system regimes. The results of these applications have implications for electricity market policy, and raise interesting questions regarding future operation and regulation of the market.

In a different approach to the blending of top-down and bottom-up models, we calibrated a Cournot market model of the Australian NEM and combined its simulated results with a top-down model for the residual variation in prices. We showed that the estimated parameters for demand and supply were in line with conventional wisdom and, provided the supply-side information used by the Cournot model accurately reflected reality, the calibrated Cournot model was able to provide accurate estimates of the underlying level of prices.

Our intention has always been to develop models that are based on sound rationale, perform well, and produce usable results, no matter how simple those models (or the ideas behind them) actually are. As a result, this research has had a highly practical focus; all the tools developed have the potential to be of use to market participants, regulators, investors, consumers and other interested parties. The models have been developed using publicly-available market data, which has the advantage that others will be able to recalibrate and apply the models in future when more data becomes available. In the course of this research, we have asked and answered many questions. However, with the tools we have developed, the door has been opened for further extensions[1] and applications to answer many more.

[1] Several of these extensions and directions for further research are identified in Appendix J. Some have been explored thoroughly, while others are still at the conceptual stage. The extensions include:

- Modelling higher frequency prices
- Using a non-symmetric distribution for the stochastic component of the price model
- Adding other physical information to the NZEM price and release models
- Estimating the MWV with piecewise linear storage envelopes
- Constructing a two-island price and release model for the NZEM
- Using the MWV estimation methodology in a Cournot model for the NZEM

REFERENCES

Aires, J.C.O., Pereira, M.V.F., Lima M.C., Barroso, L.A., & Lino, P. (2002). The effectiveness of demand elasticity in competitive hydrothermal systems (Session 15, Paper 1), Fourteenth Power Systems Computation Conference (PSCC; June 24-28, 2002), Sevilla, Spain.

Anderssen, B., & Bergmen, L. (1995). Market structure and the price of electricity: An ex ante analysis of the deregulated Swedish electricity market. *The Energy Journal, 16*(2): 97-109.

Anderson, E.J., & Philpott A.B. (2002). Using supply functions for offering generation into an electricity market. *Operations Research, 50*(3): 477-489.

Arellano, S. (2002). Diagnosing market power in Chile's electricity industry. Documento de Trabajo, No. 214, Instituto de Economía, Pontificia Universidad Católica de Chile (IE-PUC).

Atkins, F.J., & Chen, J. (2002). Some statistical properties of deregulated electricity prices in Alberta. Discussion Paper 2002-06, Department of Economics, University of Calgary, Canada.

Audet, N., Heiskanen, P., Keppo, J., & Vehviläinen, I. (2004). Modeling electricity forward curve dynamics in the Nordic market. In D.W. Bunn (Ed.), *Modelling prices in competitive electricity markets* (p. 251-266). Chichester: John Wiley & Sons, Ltd.

Baíllo, A., Ventosa, M., Ramos, A., Rivier, M., & Canseco, A. (2001). Strategic unit commitment for generation companies in deregulated electricity markets. In B.F. Hobbs, M.H. Rothkopf, R.P. O'Neill, & H.–P. Chao (Eds.). *The next generation of electric power unit commitment models* (p. 227-248). Boston: Kluwer Academic Publishers.

Barlow, M.T. (2002). A diffusion model for electricity prices. *Mathematical Finance, 12*(4), 287-298.

Batstone, S.R.J. (2003). Aspects of risk management in deregulated electricity markets: Storage, market power and long-term contracts. Ph.D. Thesis, Department of Management, University of Canterbury, New Zealand.

Bessembinder, H., & Lemmon, M.L. (2002). Equilibrium pricing and optimal hedging in electricity forward markets. *The Journal of Finance, LVII*(3), 1347-1382.

Bierbrauer, M., Trück, S., & Weron, R. (2004). *Modelling electricity prices with regime switching models*. Lecture Notes (pp. 859-867), Computer Science 3039, University of Karlsruhe, Germany.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics, 31*, 307-327.

Borenstein S., Bushnell J., & Wolak F. (2000a). Diagnosing market power in California's deregulated wholesale electricity market. Program on Workable Energy

Regulation (POWER) Working Paper (PWP-064), University of California, Berkeley.

Borenstein, S., & Bushnell, J. (1999). An empirical analysis of the potential for market power in California's electricity industry. *The Journal of Industrial Economics*, *47*(3), 285-323.

Borovkova, S., & Permana, F.J. (2004). Modelling electricity prices by the potential jump-diffusion. Paper presented at Stochastic Finance 2004, ISEG, Lisbon.

Boshier, J.F., Manning, G.B., & Read, E.G. (1983). Scheduling releases from New Zealand's hydro-reservoirs. *Transactions of the Institution of Professional Engineers New Zealand (EMCh), 10*(2), p. 33.

Botterud, A., Bhattacharyya, A.K., & Ilic, M. (2002). Futures and spot prices - An analysis of the Scandinavian electricity market. Proceedings of the Thirty-fourth Annual North American Power Symposium (NAPS; 14-15 October), Tempe, Arizona.

Bunn, D. (2004). Structural and behavioural foundations of competitive electricity prices. In D.W. Bunn (Ed.), *Modelling prices in competitive electricity markets* (p. 1-17). Chichester: John Wiley & Sons, Ltd.

Bunn, D.W., & Karakatsani, N. (2003). *Forecasting Electricity Prices*. EMG working paper, London Business School.

Burger, M., Klar, B., Müller, A., & Schindlmayr, G. (2004). A spot market model for pricing derivatives in electricity markets. *Quantitative Finance, 4*(1), 109-122.

Bushnell, J.B. (2002). A mixed complementarity model of hydrothermal electricity competition in the western United States. *Operations Research, 51*(1), 80-93.

Bushnell, J.B., & Saravia, C. (2002). An empirical assessment of the competitiveness of the New England electricity market. Center for the Study of Energy Markets Working Paper, 101 (CSEM WP-101), *University of California Energy Institute (UCEI)*, Berkeley.

Bushnell, J.B. (2003). Looking for trouble: Competition policy in the U.S. electricity industry. Center for the Study of Energy Markets Working Paper, 109 (CSEM WP-109), *University of California Energy Institute (UCEI)*, Berkeley.

Bushnell, J.B., & Mansur, E.T. (2005). Consumption under noisy price signals: A study of electricity retail rates in San Diego. *The Journal of Industrial Economics, 53*(4), 493-514.

Bushnell, J.B., Mansur, E., & Saravia, C. (2005). Vertical relationships, market structure, and competition: An analysis of U.S. restructured electricity markets. Center for the Study of Energy Markets Working Paper, 126 (CSEM WP-126), *University of California Energy Institute (UCEI)*, Berkeley.

Cartea, A., & Figueroa, M.G. (2005). Pricing in electricity markets: A mean reverting jump diffusion model with seasonality. *Applied Mathematical Finance, 12*(4), 313-335.

Chan, K.F., & Gray, P. (2005). A new jump-diffusion model for electricity spot prices with extreme jumps. Proceedings of the Fifteenth Meeting of the New Zealand Econometric Study Group (NZESG; 19-20 August), Auckland University of Technology, New Zealand.

Chattopadhyay, D. (2004). Multicommodity spatial cournot model for generator bidding analysis. *IEEE Transactions on Power Systems, 19*(1), 267-275.

Church, J.R. & Ware, R. (2000). *Industrial Organization: A strategic approach*. Singapore: McGraw-Hill Book Co.

CRA International (Asia-Pacific) Ltd. (2003). T-CONE: A tool to aid bidding strategy based on Cournot-Nash equilibrium. Confidential Report, November.

Daellenbach, H.G. (1994). *Systems and decision making: A management science approach*. New York: John Wiley & Sons, Inc.

Daniels, C. (2005, 24 January). Power shortages consigned to dark ages. *The New Zealand Herald*, New Zealand.

Davison, M., Anderson, C.L., Marcus, B., & Anderson, K. (2002). Development of a hybrid model for electrical power spot prices. *IEEE Transactions on Power Systems 17*(2), 257-264.

de Jong, C., & Huisman, R. (2002). Option formulas for mean-reverting power process with spikes. Working Paper, Energy Global, Netherlands; ERIM Report Series Reference No. ERS-2002-96-F&A.

de Jong, C. (2005). The nature of power spikes: A regime-switch approach. ERIM Report Series Reference No. ERS-2005-052-F&A.

de Lange, A.J.P., Schavemaker, P.H., & van der Sluis, L. (2002). Electricity prices: Stochastic or deterministic? Working Paper, ADL Consultancy, Netherlands.

Deng, S. (2000a). Pricing electricity derivatives under alternative stochastic spot price models. Proceedings of the Thirty-Third Hawaii International Conference on System Sciences (January), Hawaii.

Deng, S. (2000b). Stochastic models of energy commodity prices and their applications: mean-reversion with jumps and spikes. Program on Workable Energy Regulation (POWER) Working Paper (PWP-073), University of California Energy Institute (UCEI), Berkeley.

Deng, S. Jiang, W., & Zhendong, X. (2002). Alternative statistical specifications of commodity price distribution with fat tail. *AMO – Advanced Modeling and Optimization, 4*(2), 1-8.

Duffie, D., Gray, S., & Hoang, P. (1999). Volatility in energy prices. In R. Jameson (Ed.), *Managing energy price risk* (2nd ed.). London: Risk Publications.

Engle, R. (2001). GARCH 101: The use of ARCH/GARCH models in applied econometrics. *Journal of Economic Perspectives, 15*(4), 157–168.

Engle, R.F. (2002). New frontiers for ARCH models. *Journal of Applied Econometrics* , *17*(5), 425-446.

Escribano, A., Peña, J.I., & Villaplana, P. (2002). Modelling electricity prices: International evidence. Working Paper 02-27, Economic Series 08, Universidad Carlos III de Madrid.

Ethier, R., & Mount, T. (1998). Estimating the volatility of spot prices in restructured electricity markets and the implications for option values. PSerc Working Paper, Cornell University.

Evans, J., & Green, R. (2005). Why did British electricity prices fall after 1998? Working Paper, University of Surrey and University of Birmingham, Great Britain.

Eydeland, A., & Wolyniec, K. (2003). Energy and power risk management: New developments in modelling, pricing and hedging (1st Ed.). New Jersey: John Wiley and Sons, Inc.

Fleten, S.-E., & Lemming, J. (2003). Constructing forward price curves in electricity markets. *Energy Economics 25*(5), 409-424.

Frame, R., & Joskow, P. (1998). Testimony in State of New Jersey Board of Public Utilities Docket No. EX94120585Y and EO97070463, 1998.

Garcia, A., & Arbelaez, L (2002). Market power analysis for the Colombian electricity wholesale market. *Energy Economics, 24*(3), 217-229.

Garcia, A., Campos-Nañez, E., & Reitzes, J. (2005). Dynamic pricing and learning in electricity markets. *Operations Research, 53*(2), 231-241.

Geman, H., & Roncoroni, A. (2006). Understanding the fine structure of electricity prices. *Journal of Business, 79*(3), 1225-1262.

Gjolberg, O., & Johnsen, T. (2001). Electricity futures: Inventories and price relationships at Nord Pool. Proceedings of the Twelfth Annual Asia-Pacific Futures Research Symposium (Discussion Paper #D-16/2001), Bangkok, Thailand.

González, A.M., San Roque, A.M., & García-González, J. (2005) Modelling and forecasting electricity prices with input/output hidden markov models. *IEEE Transactions on Power Systems*, *20*(1), 13-24.

Gorman, P. (2005, 15 November). Low lakes spark fear of power crisis. *The Press*, New Zealand.

Gorman, P., & Steere, M. (2005, 6 December). NZ resorts to Plan B over power, *The Press*, New Zealand.

Goto, M., & Karolyi, G.A. (2004). Understanding electricity price volatility within and across markets. Dice Center Working Paper (No. 2004-12), Ohio State University, Ohio.

Green, R.J., & Newberry, D.M. (1992). Competition in the British electricity spot market. *Journal of Political Economy, 100*(5), 929-953.

Green, R.J. (2003). Electricity markets: Challenges for economic research. Proceedings of the Research Symposium on European Electricity Markets (26 September), The Hague, Netherlands.

Green, R.J. (2004). Did English generators play Cournot? Capacity withholding in the Electricity Pool. CMI Working Paper 41, University of Cambridge, Great Britain.

Guirguis, H.S., & Felder, F.A. (2004). Further advances in forecasting day-ahead electricity prices using time series models. *KIEE International Transactions on PE, 4-A*(3), 159-166.

Guthrie, G.A., & Videbeck, S. (2002a). The Marginal Cost of Electricity: What is water worth? *New Zealand Institute for the Study of Competition and Regulation Inc. (ISCR) Competition and Regulation Times, 8*, 1&10.

Guthrie G.A., & Videbeck, S. (2002b). High frequency electricity spot price dynamics: An intra-day markets approach. Working Paper, New Zealand Institute for the Study of Competition and Regulation Inc. (ISCR).

Guthrie G.A., & Videbeck, S. (2004). Electricity spot price dynamics: Beyond financial models. Working Paper, New Zealand Institute for the Study of Competition and Regulation Inc.. (ISCR) and Victoria University of Wellington, New Zealand.

Haldrup, N., & Neilsen, M. (2004). A regime switching long memory model for electricity prices. Working Paper, University of Aarhus, Denmark.

Harte, D., & Thomson, P. (2004). Stochastic models for hydro catchment inflows: An exploratory analysis. Report for the New Zealand Electricity Commission, Statistics Research Associates Ltd., New Zealand.

Heavy rains help ease power prices (2006, 19 January). *www.xtramsn.co.nz*.

Huisman R., & Mahieu, R. (2003). Regime jumps in electricity prices. *Energy Economics, 25*(5), 425-434.

Johnsen, T.A. (2001). Demand, generation and price in the Norwegian market for electric power. *Energy Economics, 23*(3), 227-251.

Johnson, B., & Barz, G. (1999). Selecting stochastic processes for modelling electricity prices. In *Energy modelling and the management of uncertainty*. London: Risk Publications.

Kaminski, V. (1997). The challenge of pricing and risk managing electricity derivatives. In P. Barber (Ed.), *The U.S. power market: Restructuring and risk management* (p. 149-171). London: Risk Publications.

Karakatsani, N.V., & Bunn, D.W. (2004). Modelling the volatility of spot electricity prices. Working Paper, London Business School.

Knittel, C.R., & Roberts, M.R. (2005). An empirical examination of restructured electricity prices. *Energy Economics, 2* (5), 791-817.

Koopman, S.J., Ooms, M., & Carnero, M.A. (2005). Periodic seasonal Reg-ARFIMA-GARCH models for daily electricity spot prices. Tinbergen Institute Discussion Papers 05-091/4, Tinbergen Institute, The Netherlands.

Langmore, M., & Dufty, G. (2004). Domestic electricity demand elasticities: Issues for the Victorian Energy Market. St Vincent de Paul Society Victoria Inc., Australia.

Laufer, F., & Morel-Seytoux, H.J. (1979). Optimal weekly releases from a seasonal reservoir, I. Deterministic future. *Water Resources Research, 15*(2), 383-398.

Leonard, G.R., Reitzes, J.D., Schumacher, A.D., & Bohn, J.G. (2002). Regional interactions in electricity prices in the Eastern United States. In M.A. Crew & J.C. Schuh (Eds.), *Markets, pricing and deregulation of utilities* (p. 109-142). Boston: Kluwer Academic Publishers.

Leyland, B. (2004, 3 September). Power planners play with chilling future, *The New Zealand Herald*, New Zealand.

Li. Y., & Flynn, P.C. (2004a). Deregulated power prices: Comparison of diurnal patterns. *Energy Policy, 32*, 657–672.

Li, Y., & Flynn, P.C. (2004b). Deregulated power prices: Comparison of volatility. *Energy Policy, 32*, 1591–1601.

Longstaff, F.A., & Wang, A.W. (2004). Electricity forward prices: A high frequency empirical analysis. *The Journal of Finance, 59*(4), 1877- 1900.

Lora, A.T., Santos, J.C.R., Santos, J.M.R, Ramos, J.L.M., & Expósito, A.G. (2002). Electricity market price forecasting: Neural networks versus weighted-distance k nearest neighbours. Proceedings of the Thirteenth International Conference on Database and Expert Systems Applications (2-6 September). Aix-en-Provence, France.

Lucia, J.J., & Schwartz, E.S. (2002). Electricity prices and power derivatives: Evidence from the Nordic Power Exchange. *Review of Derivatives Research, 5*(1), 5-50.

MacDonald, A. (2004, 10 September). Meridian calls for water resources to be traded as commodity, *www.stuff.co.nz*.

Makridakas, S.G., Wheelwright, S.C., & Hyndman, R.J. (1998). *Forecasting: Methods and applications* (3rd Ed.). New York: John Wiley & Sons, Inc.

Mason, G. (2002). A comparison of restructured electricity pools, the key characteristics of NZEM spot prices, and the estimation of a two factor mean reverting process. Masters of Commerce and Administration (MCA) Thesis, Victoria University of Wellington, New Zealand.

Miller, I., & Miller, M. (2004). *John E. Freund's Mathematical Statistics with Applications* (7th Ed.). London: Pearson Education.

Mount, T.D., Ning, Y., & Cai, X (2006). Predicting price spikes in electricity markets using a regime-switching model with time-varying parameters. *Energy Economics, 28*(1), 62-80.

Natarajan, H. (2003). Modeling electricity spot price behavior: Seasonality, mean reversion, heteroscedasticity and jumps. Working Paper, Indian Institute of Management Bangalore, India.

National Institute of Economic and Industry Research (2005). The price elasticity of demand for electricity in NEM regions. Prepared for the National Electricity Market Management Company (June 2005), Australia.

Neuhoff, K., Barquin, J., Boots, M.G., Ehrenmann, A., Hobbs, B.F., Rijkers, F.A.M., & Vázquez, M. (2005). Network-constrained Cournot models of liberalized electricity markets: The devil is in the details. *Energy Economics*, *27*(3), 495-525.

Nogales, F.J., Contreras, J., Conejo, A.J., & Espínola, R. (2002). Forecasting next-day electricity prices by time series models. *IEEE Transactions on Power Systems, 17*(2), 342-348.

New Zealand Press Association (2004, 18 August). Big chill causes record power usage. *www.stuff.co.nz*.

Pankratz, A. (1991). *Forecasting with dynamic regression models*. New York: John Wiley & Sons, Inc.

Pirrong, C. & Jermakyan, M. (1999). Valuing power and weather derivatives on a mesh using finite difference methods. In R. Jameson (Ed.), *Energy modelling and the management of uncertainty*. London: Risk Publications.

Pirrong, C., & Jermakyan, M. (2001). The price of power: The valuation of power and weather derivatives. Working paper, Oklahoma State University.

Popova, J. (2004). Spatial pattern in modeling electricity prices: Evidence from the PJM market. Working Paper, West Virginia University.

Power prices rise amid cold snap. (2004, 23 December). *www.tvnz.co.nz*.

Puller, S.L. (2007). Pricing and firm conduct in California's Deregulated Electricity Market. *Review of Economics and Statistics*, *89*(1), 75-87.

Read, E.G. (1979). Optimal operation of power systems. Ph.D. Thesis, Department of Economics, University of Canterbury, New Zealand.

Read, E.G. (1984). Deterministic reservoir operation – An application of the economic principles. Report ER 4006, Ministry of Energy, New Zealand.

Read, E.G., Culy, J.G., Halliburton, T.S., & Winter, N.L. (1988). A simulation model for electricity planning in New Zealand. In G. Rend (Ed.), *Operational Research '87*. Proceedings of the Eleventh IFORS International Conference, Buenos Aires, Argentina, 10-14 August 1987 (p. 493-507). Amsterdam: North-Holland.

Read, E.G., & Boshier, J.F. (1989). Biases in stochastic reservoir scheduling models . In A.O. Esogbue (Ed.), *Dynamic programming for optimal water resources system management* (p. 386-398). New York: Prentice Hall.

Robinson, T.A. (2000). Electricity pool prices: A case study in nonlinear time-series modelling. *Applied Economics, 32*(5), 527-532.

Said, S.E., & Dickey, D.A. (1984). Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika, 71*(3), 599–607.

Sansom, D.C., & Saha, T.K. (1999). Neural networks for forecasting electricity pool prices in a deregulated electricity supply industry. Working Paper, University of Queensland, Australia.

Scott, T.J., & Read E.G. (1996). Modelling hydro reservoir operation in a deregulated electricity market. *International Transactions in Operational Research, 3*(3-4), 209-221.

Scott, T.J. (1998). Hydro reservoir management for an electricity market with long-term contracts. Ph.D. Thesis, Department of Management, University of Canterbury, New Zealand.

Skantze, P., Gubina, A., & Ilic, M. (2000). Bid-based stochastic model for electricity prices: The impact of fundamental drivers on market dynamics. Working Paper (MIT EL 00-004) Energy Laboratory, Massachusetts Institute of Technology.

Steeman, M. (2004, 24 November). Hodgson rules out price control, *The Dominion Post*, New Zealand.

Vehviläinen, I., & Pyykkönen, T. (2005). Stochastic factor model for electricity spot price – the case of the Nordic market. *Energy Economics, 27*(2), 351-367.

Videbeck, S. (2004). Electricity spot price behavior: Insights into the NZEM. Masters Thesis, Department of Economics, Victoria University of Wellington, New Zealand.

Villaplana, P. (2003). Pricing power derivatives: A two-factor jump-diffusion approach. Job Market Paper, Universidad Carlos III de Madrid, Spain.

Villaplana, P. (2005). Valuation of electricity forward contracts: The role of demand and capacity. Working Paper. Department of Economics and Business, Universitat Pompeu Fabra, Barcelona, Spain.

Vucetic, S., Tomsovic, K., & Obradovic, Z. (2001). Discovering price-load relationships in California's electricity market. *IEEE Transactions on Power Systems, 16*(2), 280–286.

Weron R., & Misiorek, A. (2005). Forecasting spot electricity prices with time series models. Proceedings of the International Conference, The European Electricity Market EEM-05 (10-12 May), Lodz, Poland.

Westergaard, E., Mullon, G., Sise, G., & McCord, A. (1996). Development of TransPower's wholesale pool simulation model. Working paper, Transpower New Zealand Ltd.

Willems, B. (2005). Cournot competition, financial option markets and efficiency. Center for the Study of Energy Markets Working Paper, 139 (CSEM WP-139), *University of California Energy Institute (UCEI)*, Berkeley.

Wilson, D.I., & Cheng, B.C.M. (2004). Predicting the unpredictable: Is the electrical spot price chaotic? Working Paper, Auckland University of Technology, New Zealand. In C. Boyle & R. MacDowall (Eds.), *International Conference on Sustainability Engineering*, Auckland, New Zealand.

Wolak, F.A. (1997). Market design and price behavior in restructured electricity markets: An international comparison. Working Paper, Department of Economics, Stanford University, California.

Wolak, F.A. (2000). An empirical analysis of the impact of hedge contracts on bidding behavior in a competitive electricity market. *International Economic Journal*, *14*(2), 1-39.

Wolak, F.A. (2003a). Measuring unilateral market power in wholesale electricity markets: the California market 1998 to 2000. *American Economic Review, 93*(2), 425-430.

Wolak, F.A. (2003b) Identification and estimation of cost functions using observed bid data: an application to electricity markets. In M. Dewatripont, L.P. Hansen, & S.J.

Turnovsky (Eds.), *Advances in economics and econometrics: Theory and applications*, *Vol. II* (p. 115-149). New York: Cambridge University Press.

Wolak, F.A., & Patrick, R.H. (1997). The impact of market rules and market structure on the price determination process in the England and Wales electricity market. Program on Workable Energy Regulation (POWER) Working Paper (PWP-047), University of California Energy Institute (UCEI), Berkeley.

Wu, G. (2001). The determinants of asymmetric volatility. *The Review of Financial Studies, 14*(3), 837-859.

Yang, M. (1995). Dual dynamic programming for reservoir management with correlated inflows. Ph.D. Thesis, Department of Management, University Of Canterbury, New Zealand.

Appendix A

ELECTRICITY WHOLESALE MARKET OPERATION

In order to understand how a deregulated wholesale electricity market operates, it is necessary first to explain the most fundamental part of economic theory – the theory of demand and supply.

Imagining the theory behind a simple, well-behaved one-good market on its own will aid in understanding how supply and demand in an electricity market is balanced. In most markets, the higher the price of a good rises, the more of the good suppliers will be willing to sell. Conversely, the higher the price of a good is, the less of it consumers will be willing to purchase. The opposite is also true – if the price of the good falls, consumers will want to purchase more of the good but suppliers will be willing to sell less. These effects are represented graphically in Figure A.1 below by an upward-sloping supply curve and downward-sloping demand curve respectively.

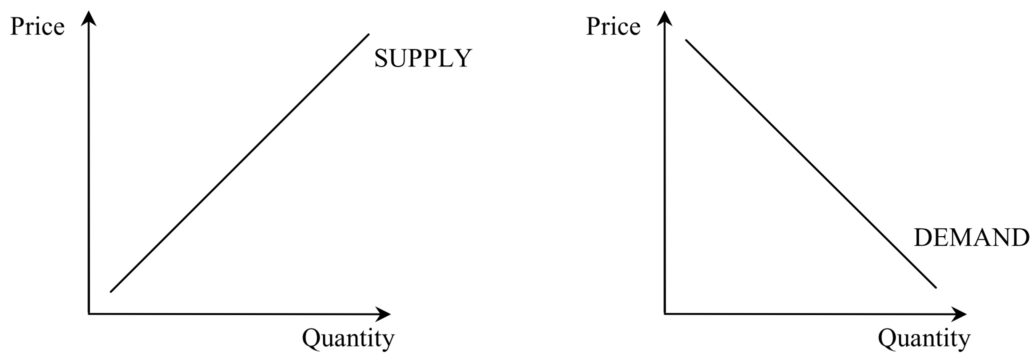

Figure A.1: Supply and demand curves in a basic market

From the two graphs above, the higher the price is, the greater is the supply and the lower is the demand, and vice versa. The actual price of the good is determined where the supply and the demand curves intersect. The point of intersection gives the price and quantity for which the exact amount of the good supplied will be consumed; no one who demands some of the good at that price will go away empty-handed, and no supplier will be left with any units of the good unsold. When this is the case it is said that the market has been cleared, and the intersection point yields the market-clearing price and the market-clearing quantity. On the graph below of the intersection of demand and supply, the market clearing price is denoted by p* and the market-clearing quantity is q*:

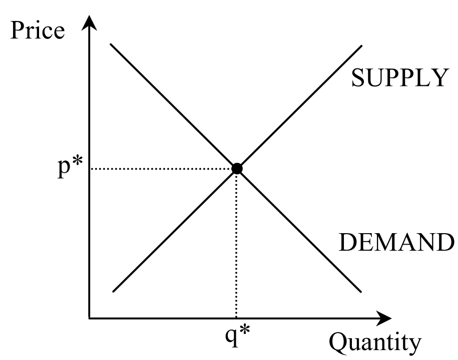

Figure A.2: Market-clearing price and quantity in a basic market

The wholesale electricity market is cleared using exactly the same principles, and can be explained in the same way. The suppliers of electricity to the wholesale spot market are the generating companies, who own and operate power generating plants. The consumers of electricity are retail power companies (who effectively on-sell the power to households and businesses) and some large power users. The market is cleared to determine which generators will be required to produce power (be "dispatched"), what their dispatch quantities will be, and what the spot price of power will be.

The supply curve in a wholesale electricity market is an accumulation of each of the generating companies' individual supply curves. These are referred to as their "offer" curves or "offer stacks", as they give the prices at which each company is willing or offering to supply specific quantities of power to the market. For example, in a perfectly competitive market, in which each company offers to generate electricity at the marginal cost of generation, electricity is offered in blocks corresponding to capacities of generating units, at the cost at which that those units operate. If a company (Company A) owns one unit (Unit 1) that can generate 10 megawatts (MW) at a cost of $10 per megawatt per hour (MWh), and another unit (Unit 2) that can generate 5 MW at $20/MWh, then they might offer their generation to the market in those two steps.

The accumulated (market) offer stack is upward-sloping and step-shaped, to reflect the fact that different technologies are used to generate power for different blocks, each with a different cost of generation. Simplistically, if a generating company wants to supply a certain amount of power, they will first turn on their cheapest generator to run (this would be Unit 1 for Company A). If the company wants to generate more power than the capacity of their cheapest generator, they will turn on their next-cheapest generator (Unit 2), and so on. A generating company will only be willing to run a unit if the market price of power is at least as great as the price at which it is able to produce. If it is not, then they cannot generate electricity profitably at the market price.

In reality, the offer stack is not necessarily a true reflection of the marginal cost of generation for each generating unit. Some of the larger generators (such as conventional thermal generators) have high start-up and shut-down costs but are cheap to run, and thus once committed will be offered at a very low price to ensure that they will be dispatched. Companies also have the incentive to ensure that they produce at least as much power as they are contracted to supply (their “contract level”), to limit the costs (and risks) they would otherwise have to bear through purchasing electricity from the spot market to meet their contractual obligations. As a result, they will often offer to generate up to their contract level at a very low price. In the situation where water for hydro generators is scarce, companies will also be very unwilling to dispatch these generators for anything other than an extremely high price, so as not to compromise their ability to operate profitably in a future period.

A further point to note is that at times when the system capacity is stretched and generators know that their more expensive generation units are likely to be dispatched, they can increase the price at which they offer these units, to increase the market price. This introduces the issue of gaming by the players in the market – the higher the price at which every company in the market offers to generate, the higher the market-clearing price will be, but obviously if one player undercuts another’s offer, then they will generate more power and capture some of their rivals’ potential profits. All of the factors mentioned above lead to the market offer stack being virtually flat at low levels of power, and much steeper at the top end.

In the short term, the demand curve for electricity in an electricity market is almost vertical, indicating near perfect inelasticity of demand. This is partly because, through the use of retail contracts, many users of electricity (i.e. households and small businesses) are effectively not exposed to fluctuations in the wholesale price of power, and their decisions on usage are not influenced by the price which their retail company is paying for it. But it is also the case that, once committed to particular technologies, consumers may have very little ability to adjust their electricity usage in the short term. In the longer term, however, their decisions will be more sensitive to the price as retail companies pass changes in the spot price through to their customers. Consumers who purchase their power directly from the spot market are more responsive to increases in the price and may make decisions to shut down production when the price gets too high.

For simplicity, the following demand and supply diagram ignores demand-side bids and includes a near-vertical short-term demand curve to show how the electricity market is cleared. Further, the diagram assumes that the market has only one node, or equivalently that transmission of power is unconstrained and lossless. Thus the market-clearing supply and demand curves can be represented in simple form as shown in Figure A.3 below.

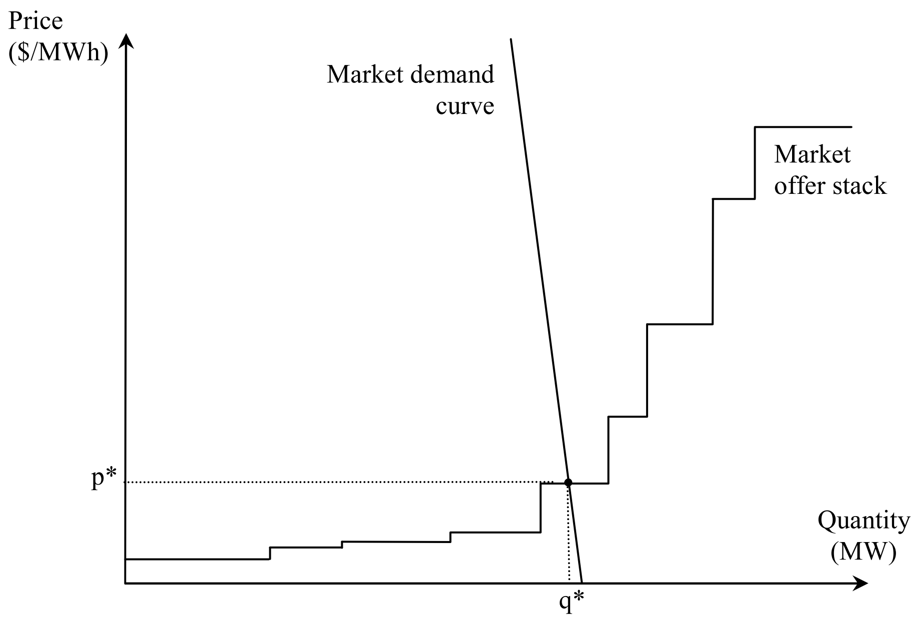

Figure A.3: Supply and demand curves and market-clearing point in an electricity market

In the situation illustrated in Figure A.3 above, the spot price will be p* and the total dispatch quantity of power generated will be q*. Each generating unit offered for a price below p* is fully dispatched, and the unit corresponding to the offer step intersected by the demand curve is only partially dispatched, up to the required amount. This generation unit is referred to as the marginal generator, as it is on the margin between being fully dispatched or not dispatched at all. Every generating unit which is offered at a price above p* is not dispatched at all.

Obviously, the amount of power required by end users changes every micro-second as consumer use varies. In order that no customers ever experience a shortage of power (i.e. the light bulb always turns on when you flick the switch), demand and supply must be constantly balanced. This balance is also required to maintain the frequency and stability of the transmission network. The market is cleared (as in Figure A.3) at fixed intervals throughout every day – this happens every half-hour in New Zealand; in Australia it is every five minutes. This means that each genco submits a different offer curve for every half hour period, and the market is cleared with a different p* and q* every time.

In between each market-clearance, the shortfall between demand and supply is balanced using reserve generation, which is dealt with in another market (the reserve market). The reserve generators are able to increase or decrease generation rapidly to maintain the balance between supply and demand. In some countries, such as New Zealand, the energy and reserve markets are cleared simultaneously, while other markets (i.e. Australia) are energy only. The research in this thesis deals with energy markets (and the energy sides of simultaneously-cleared markets) only[1], which are cleared using the principles and methods described above.

[1] In markets such as New Zealand's, spot prices are calculated several times for a single market clearance. Prices are forecasted up to 35 hours ahead of market clearance (forecast prices), four hours before market clearance (dispatch prices), at the time of market clearance (real-time prices) and the day after market clearance (final prices). Unless stated otherwise, all prices analysed in this thesis are final spot prices.

Appendix B

MAXIMUM LIKELIHOOD ESTIMATION

One of the common methods of parameter estimation employed in mathematical statistics is the method of Maximum Likelihood Estimation (MLE). In essence, given a set of observations and a particular specification of a model for generating those observations, MLE finds the set of parameters for that model that are most *likely* to have (or have the highest probability of having) generated the particular set of observations observed. As with any parameter estimation technique, the parameters estimated by MLE are unique to the model specified and the observations the model is being fitted to.

As defined by Miller and Miller (2004):

> "Thus, the essential feature of the method of MLE is that we look at the sample values and then choose as our estimates of the unknown parameters the values for which the probability or probability density of

getting the sample values is a maximum … In the discrete case, if the observed sample values are $x_1, x_2, \ldots, x_n$, the probability of getting them is

$$P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n) = f(x_1, x_2, \ldots, x_n; \theta)$$

which is just the joint probability distribution of the random variables $X_1, X_2, \ldots, X_n$ at $X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n$. Since the sample values have been observed and are therefore fixed numbers, we regard $f(x_1, x_2, \ldots, x_n; \theta)$ as a value of a function of θ, and we refer to this function as the likelihood function."

B.1 The likelihood function

The likelihood function in the model of Escribano et al. (2002), among others, is derived from the probability density function (PDF) for a mixed Poisson-Gaussian distribution. The PDF for a normally distributed (or Gaussian) variable x with mean μ and variance σ^2 is:

$$f\left[x \mid \mu, \sigma^2\right] = \exp\left[\frac{-(x-\mu)^2}{2\sigma^2}\right]\frac{1}{\sqrt{2\pi\sigma^2}}$$

Consider the case where a variable x has a certain probability λ of coming from one normal distribution, and a probability $(1-\lambda)$ of coming from another. In that case, its probability density function would be:

$$\begin{aligned} f\left[x \mid \lambda, \mu_1, \sigma_1^{\,2}, \mu_2, \sigma_2^{\,2}\right] \quad = \quad & \lambda.\exp\left[\frac{-(x-\mu_1)^2}{2\sigma_1^{\,2}}\right]\frac{1}{\sqrt{2\pi\sigma_1^{\,2}}} \\ & + (1-\lambda).\exp\left[\frac{-(x-\mu_2)^2}{2\sigma_2^{\,2}}\right]\frac{1}{\sqrt{2\pi\sigma_2^{\,2}}} \end{aligned}$$

Given the discrete-time price model of Escribano et al. (2002) specified in Chapter 3, the following mixed Poisson-Gaussian density function can be derived for the price on day *t*, given the price on day *t*-1:

$$f[p_t \mid p_{t-1}] = \lambda_t.\exp\left[\frac{-\left(p_t - f(t) - \phi.p_{t-1} - \mu_t\right)^2}{2(h_t + \sigma_t^{\ 2})}\right]\frac{1}{\sqrt{2\pi(h_t + \sigma_t^{\ 2})}}$$
$$+ (1-\lambda_t).\exp\left[\frac{-\left(p_t - f(t) - \phi.p_{t-1}\right)^2}{2h_t}\right]\frac{1}{\sqrt{2\pi h_t}}$$

The likelihood function for the entire sample period can then be approximated by the following:

$$\mathrm{L}(\Theta) = \prod_{t=1}^{T} f[p_t \mid p_{t-1}]$$

Conveniently, the set of parameter estimates Θ that maximises that function will also maximise the natural log of L(Θ). Hence it is conventional to maximise the log of the likelihood function, which produces the *log-likelihood*. The greater the log-likelihood for a particular set of data, the greater the fit of that model to the data.

The set of parameter values that maximises this function is then found using traditional non-linear optimisation techniques involving gradient-search methods. Our parameter estimation was conducted in GAUSS 6.0.

Appendix C

A TRENDED LOWER STORAGE ENVELOPE

One extension to the model for NZEM spot prices proposed in Chapter 5 involves a linear trend in the lower storage envelope. The estimated parameters of this extended model are listed in this appendix.

MWV model including linearly-trended lower storage envelope		
1 Aug 1999 - 30 June 2003		
	Coefficient	t-stat
δ	0.22214	0.12
c_{COLD}	22.6347	897.08
w_{COLD}	406.5253	372551.09
x_{COLD}	-0.05475	-0.28
y_{COLD}	11.11476	68.53

z_{COLD}	1.6218	63.01
c_{WARM}	11.4171	402.25
w_{WARM}	110.4359	26589.81
x_{WARM}	-0.1089	-1.60
y_{WARM}	7.2928	72.11
z_{WARM}	1.2338	72.82
WEEKDAY	3.1302	33.38
θ	0.8355	42.58
ω	3.2379	51.46
α	0.3543	28.21
β	0.5475	31.37
λ_{COLD}	8.2419%	5.11
μ_{COLD}	19.5927	2874.51
σ^2_{COLD}	249.8912	1425049.12
λ_{WARM}	5.2208%	2.36
μ_{WARM}	12.9432	1948.51
σ^2_{WARM}	382.6298	1782660.20
Log-Likelihood	-5085	
SIC	10331	

Table C.1: Estimated parameters from application of EPV model including two seasonal MWV functions and two seasonal jump distributions and a linearly-trended lower storage envelope

As noted in Chapter 5, while the log-likelihood indicates an improved fit over the model without the trended envelope, the estimated parameter values for δ, x_{COLD} and x_{WARM} are not statistically significant.

Appendix D

MWV-BASED PRICE STOCHASTICITY

This appendix contains tables of the estimated models including MWV-based stochasticity, as discussed in Chapter 5. The parameters of the initial model, with constant stochasticity, are listed in Table D.1, and the estimated parameters of the final, adjusted model, with a mixture of constant and MWV-based stochasticity, are listed in Table D.2.

D.1 Base model before any stochastic component extensions

EPV Model including two MWV functions		
1 Aug 1999 - 30 June 2003		
	Coefficient	t-stat
c_{COLD}	22.1911	991.05
w_{COLD}	481.2378	755688.08
x_{COLD}	-0.1519	-2.46

EPV Model including one MWV function		
1 Aug 1999 - 30 June 2003		
	Coefficient	t-stat
c	8.7168	398.50
w	3752.1734	29422202.07
x	-2.104735	-188.42

y_{COLD}	7.1935	55.02
z_{COLD}	1.3351	62.33
c_{WARM}	8.8200	347.62
w_{WARM}	115.6951	28464.76
x_{WARM}	-0.3845	-19.41
y_{WARM}	2.7616	28.39
z_{WARM}	0.7824	52.00
WEEKDAY	3.0561	32.43
θ	0.8616	40.12
ω	2.8431	41.53
α	0.3364	24.84
β	0.5735	30.79
λ_{COLD}	8.1675%	5.19
μ_{COLD}	17.9615	2636.38
σ^2_{COLD}	249.4454	1372291.34
λ_{WARM}	5.5604%	2.60
μ_{WARM}	12.0966	1765.66
σ^2_{WARM}	383.0205	1707156.66
Log-Likelihood	-5095	
SIC	10342	

y	6.2795	63.14
z	0.3806	7.15
WEEKDAY	3.0137	31.61
θ	0.8873	36.98
ω	2.4751	36.28
α	0.3316	23.74
β	0.5900	30.88
λ	6.9805%	2.76
μ	13.4738	1312.24
σ^2	323.91	1056360.26
Log-Likelihood	-5103	
SIC	10301	

Table D.1: Estimated parameters from application of EPV model including two seasonal MWV functions and two seasonal jump distributions and EPV model including one MWV function and one jump distribution to daily average spot prices from the Benmore node: August 1999 – June 2003

D.2 Model including adjustments to both jump probability and jump mean

Original Price Model		
1 Aug 1999 - 30 June 2003		
	Coefficient	t-stat
c	8.7168	398.50
w	3752.1734	29422202.07
x	-2.104735	-188.42
y	6.2795	63.14
z	0.3806	7.15
WEEKDAY	3.0137	31.61
θ	0.8873	36.98
ω	2.4751	36.28
α	0.3316	23.74
β	0.5900	30.88
λ	6.9805%	2.76
μ	13.4738	1312.24
σ^2	323.91	1056360.26
Log-Likelihood	-5103	
SIC	10301	

Price Model after adjustments		
1 Aug 1999 - 30 June 2003		
	Coefficient	t-stat
c	16.4030	624.54
w	424.1784	435096.75
x	-0.1245	-130.65
y	9.5898	75.4863
z	1.2710	42.86
WEEKDAY	3.0731	32.28
θ	0.8772	36.45
ω	3.0181	46.07
α	0.3110	22.62
β	0.5706	32.55
λ'	0.0025	3.15
μ'	0.4069	98.27
σ^2	323.1307	885921.53
Log-Likelihood	-5096	
SIC	10287	

Table D.2: Estimated parameters from application of EPV model including one MWV function and one jump distribution and adjusted EPV model including one MWV function and one jump distribution to daily average spot prices from the Benmore node: August 1999 – June 2003

Appendix E

A NATIONWIDE MWV PRICE MODEL

In Chapter 6, the NZEM price model was re-estimated using the national aggregate storage level and final prices from the Haywards node, whereas previous models were estimated using prices from the Benmore node and storage from the Waitaki system. The Benmore price model uses data from August 1999 to June 2003, whereas the Haywards price model uses data from April 1999 to June 2003. Table E.1 contains the estimated parameters of the two different price models, along with their estimated t-statistics. The two sets of parameters are compared in Section E.2.

E.1 The model parameters

Waitaki/Benmore price model			New Zealand/Haywards price model		
1 August 1999 - 30 June 2003			1 April 1999 - 30 June 2003		
	Coefficient	t-stat		Coefficient	t-stat
c	16.4030	624.54	c	4.8721	220.73
w	424.1784	435096.75	w	354.9417	218129.89
x	-0.1245	-130.65	x	-0.1846	-230.68
y	9.5898	75.4863	y	36.1177	1430.26
z	1.2710	42.86	z	0.7056	12.78
WEEKDAY	3.0731	32.28	WEEKDAY	3.9153	47.91
θ	0.8772	36.45	θ	0.8931	38.32
ω	3.0181	46.07	ω	6.3574	174.8574
α	0.3110	22.62	α	0.3609	33.37
β	0.5706	32.55	β	0.5070	30.03
λ'	0.0025	3.15	λ'	0.0035	8.04
μ'	0.4069	98.27	μ'	0.3227	77.71
σ^2	323.1307	885921.53	σ^2	1451.8798	2730.53
Log-Likelihood	-5096		Log-Likelihood	-5757	
SIC	10287		SIC	11609	

Table E.1: Estimated parameters from application of the price model including one MWV function and jump process as a function of the MWV: Waitaki storage level and Benmore spot prices (left) and New Zealand storage level and Haywards spot prices (right), August 1999 – June 2003

E.2 Comparison between nationwide model and Waitaki model

As mentioned in Chapter 4, few conclusions can be drawn from the values of the individual parameters of the MWV functions. Also, comparing the parameters of the two different MWV functions is not particularly useful, due to the fact that different storage levels are used for each of the functions. While both functions are based on the RSL, one is based on the Waitaki RSL, which ranges in value from -235 to 1745 GWh in the sample period, and the other is based on the New Zealand RSL, which ranges from -580 to 2400 GWh.

Consistent with the results in Chapter 4 is the fact that the unconditional variance parameter ω for the Haywards node is twice the size of that estimated for the Benmore node, suggesting that the time series from Haywards is substantially more volatile. For the same MWV, jumps in prices are more common at the Haywards node. While the jumps at Haywards have a lower expected size, however, the variance in the size of these jumps is *much* greater than at Benmore. These results are all consistent with the inferences made in Chapter 4.

Appendix F

RELEASE MODEL ESTIMATION

Chapter 6 details the estimation of a model for the natural log of New Zealand's releases from hydro reservoirs. The process of estimation involves two main steps: firstly, determining the relevant (i.e. statistically significant) explanatory variables in a model for release, and secondly, determining an appropriate process for the residuals of the model to account for some of the variation in release that the explanatory variables could not.

The first step is undertaken using two different methods. The results of running a multiple regression of the natural log of release on all the potential explanatory variables detailed in Chapter 6 yields the regression results in Table F.1. This shows which explanatory variables have statistically significant slope coefficients in the regression equation (at the 5% level of significance), and which do not. The same table is provided in less detail in Chapter 6.

F.1 Multiple regression of the natural log of release on all potential drivers

Predictor	**Coefficient**	**Standard Error**	**T-statistic**	**P-value**	**Signif.**
Constant	3.91336	0.0925	42.3	0	✓
Storage Level	3.39E-06	9.06E-06	0.37	0.708	
Price	0.00046	0.000103	4.47	0	✓
Estimated MWV	-0.0155	0.001035	-14.97	0	✓
Ln_Inflow_F1	0.18459	0.03396	5.44	0	✓
Ln_Inflow_F2	-0.46021	0.03396	-13.55	0	✓
Ln_Inflow_F3	0.16427	0.03387	4.85	0	✓
Ln_Inflow_F4	-0.04178	0.0338	-1.24	0.217	
Ln_Inflow_F5	-0.00735	0.03356	-0.22	0.827	
Ln_Inflow_F6	0.0287	0.03185	0.9	0.368	
Ln_Inflow_F7	-0.02498	0.02062	-1.21	0.226	
Ln_Inflow	0.40461	0.03392	11.93	0	✓
Ln_Inflow_L1	-0.18177	0.03392	-5.36	0	✓
Ln_Inflow_L2	0.08279	0.03384	2.45	0.015	✓
Ln_Inflow_L3	-0.0306	0.03378	-0.91	0.365	
Ln_Inflow_L4	0.02779	0.03371	0.82	0.41	
Ln_Inflow_L5	0.01571	0.03351	0.47	0.639	
Ln_Inflow_L6	-0.0072	0.03182	-0.23	0.821	
Ln_Inflow_L7	0.03769	0.0208	1.81	0.07	
Saturday	-0.08117	0.01228	-6.61	0	✓
Sunday	-0.08828	0.01233	-7.16	0	✓
S	0.166147				
R-Sq	56.70%				
R-Sq(adj)	56.20%				

Table F.1: Estimated parameters and standard errors from a multiple regression of the natural log of release on all potentially relevant drivers of release

Stepwise regression is also used as a cross-check to the first variable selection method employed. Stepwise regression involves adding variables to or removing variables from a regression equation, depending on the proportion of variance that these variables explain (as measured by the r^2) and the statistical significance of their estimated slope parameters. It can be run either by starting with no variables and adding them one by one until no further explanatory variable could be added to the equation with a statistically significant slope parameter (forward selection), or starting with all the variables included and removing them one by one until only those with statistically significant slope parameters are left (backwards selection). Alternatively, the combined forwards-backwards selection method involves starting with no variables and adding them one by one at each step, but if the slope coefficient of any variable becomes insignificant it can be removed, with a chance of being included again in a later step.

The results of running a forward-backward selection stepwise regression of release on all the potentially relevant explanatory variables are shown in Table F.2, and are explained further in Chapter 6. These results were achieved using 5% for both the entry and exit level of significance. Interestingly, no variables were removed from the selection at any stage, while ten variables were added. The first variable added is the estimated water value, which explains 26% of the variation in release. The second variable added is today's inflow, then the inflow in two days' time, then the inflow yesterday, and so on, until no more variables can be included with statistically significant slope coefficients.

The residuals of the final regression model (i.e. the equation listed in Step 10 in Table F.2) are then examined to test if they exhibit any serial correlation. The plots of the sample autocorrelation and partial autocorrelation coefficients are shown in Figure F.1 and Figure F.2 respectively, with the red bands signifying statistical significance. If any of the sample coefficients are significant (which spikes in both plots are), then the residuals exhibit serial correlation that must be accounted for somehow. The significant spikes in the first few lags on the partial autocorrelogram and the steady decay of the coefficients in the autocorrelogram suggest a process including several AR coefficients is required to account for the serial correlation.

F.2 Stepwise regression of the natural log of release on all potential drivers to select the significant explanatory variables

Step	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Constant	**4.83**	**3.77**	**4.15**	**4.28**	**4.29**	**4.30**	**4.10**	**4.11**	**4.04**	**4.02**
Est_WV	-0.02	-0.01	-0.01	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02
T-Value	-23.37	-17.87	-20.51	-21.48	-21.67	-21.93	-18.71	-18.81	-18.92	-18.76
P-Value	0	0	0	0	0	0	0	0	0	0
Inflow		0.22	0.38	0.49	0.50	0.50	0.51	0.42	0.40	0.43
T-Value		17.59	25.79	20.23	20.56	20.99	21.64	12.51	12.07	12.23
P-Value		0	0	0	0	0	0	0	0	0
Inflow_F2			-0.25	-0.26	-0.26	-0.26	-0.27	-0.33	-0.46	-0.46
T-Value			-17.24	-18.14	-18.47	-18.74	-18.92	-15.23	-13.82	-13.69
P-Value			0	0	0	0	0	0	0	0
InflowL1				-0.13	-0.13	-0.13	-0.15	-0.12	-0.12	-0.18
T-Value				-5.82	-5.91	-6.14	-7.05	-5.50	-5.50	-5.42
P-Value				0	0	0	0	0	0	0
Sunday					-0.08	-0.09	-0.08	-0.09	-0.09	-0.09
T-Value					-5.74	-6.82	-6.38	-6.52	-6.62	-6.57
P-Value					0	0	0	0	0	0
Saturday						-0.09	-0.08	-0.08	-0.08	-0.08
T-Value						-6.47	-6.38	-6.43	-6.28	-6.22
P-Value						0	0	0	0	0
Price							0.00	0.00	0.00	0.00
T-Value							4.07	4.15	4.31	4.31
P-Value							0	0	0	0
Inflow_F1								0.13	0.19	0.18
T-Value								3.92	5.32	5.05
P-Value								0	0	0
Inflow_F3									0.11	0.11
T-Value									5.11	4.97
P-Value									0	0
InflowL2										0.05
T-Value										2.37
P-Value										0.018
S	0.22	0.20	0.19	0.18	0.18	0.18	0.18	0.18	0.18	0.17
R-Sq	26.06	38.37	48.30	49.41	50.47	51.78	52.99	53.45	54.23	54.39
R-Sq(adj)	26.01	38.29	48.20	49.28	50.31	51.59	52.74	53.18	53.93	54.07

Table F.2: Estimated parameters, t-values and p-values from a forward and backward selection stepwise regression (alpha and beta = 0.05) of the natural log of release on all potential drivers of release

F.3 Autocorrelogram and partial autocorrelogram of multiple regression model residuals

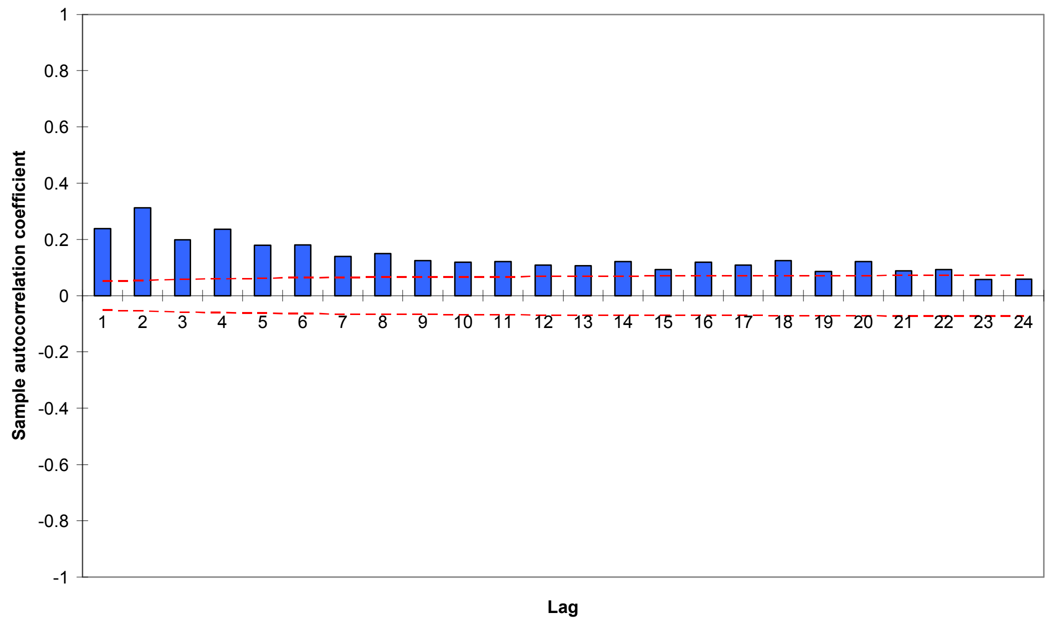

Figure F.1: Autocorrelogram of the residuals of the final regression equation listed in Table F.2

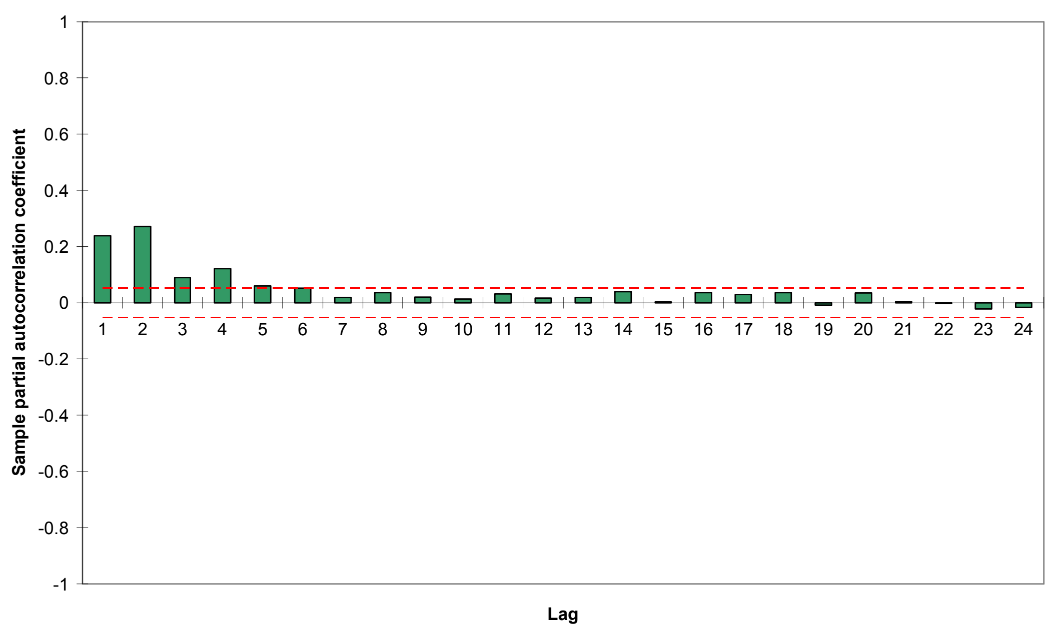

Figure F.2: Partial autocorrelogram of the residuals of the final regression equation listed in Table F.2

Several different statistics were examined in the selection of an appropriate process to account for the serial correlation in the Dynamic Regression model residuals[1]. The aim of such a fitting process is to find a process that eliminates as much of the serial correlation (or accounts for as much of the variation) in the model residuals using as parsimonious a model as possible. Including more autoregressive (AR) and moving average (MA) terms than are necessary may even introduce patterns into the residuals of a model that were not present beforehand.

The two fitting statistics used were the Akaike Information Criterion (AIC) and the Schwartz Information Criterion[2] (SIC). For both of these criteria, the smaller the value of the statistic, the better is the fit. The variance of the distribution of the errors was also examined, to ensure that it was minimized. Finally, the sample autocorrelation coefficients of the residuals of the combined model may be compared with Ljung-Box Chi-square statistics[3] to ensure that there are no statistically significant coefficients, and thus no serial correlation left in the model residuals. P-values greater than 0.05 in the bottom section of Table F.3 indicate the ARMA models for which the model residuals do not exhibit any significant serial correlation.

The final model selected for the residuals of the Dynamic Regression model is an ARMA (3,1) model. This model yielded the best fit according to the SIC and close to the best according to the AIC, as well as having a low variance and no evidence of serial correlation in the residuals.

---

[1] The autocorrelogram on the preceding page indicates that the series of residuals does not need to be differenced before an ARMA model is fitted, as the correlations at all lags are much less than one.

[2] Both of these statistics are commonly used in the assessment of the fit of time series models. Both are functions of the log-likelihood, weighted by the number of parameters in the model.

[3] This statistic was proposed by Ljung and Box (1978, as cited in Makridakis et al, 1998, p. 319) to test whether or not a residual process can be classed as white noise (i.e. there are no patterns in the residuals). The resulting statistic is compared to a Chi-square distribution with the relevant degrees of freedom; if the statistic is relatively small enough (and the corresponding p-value is greater than, say, 0.05), the null hypothesis of white noise cannot be rejected.

F.4 Selection of an ARMA process for the model residuals

Variance Estimate

		MA				
		0	1	2	3	4
AR	0	0.02786	0.02684	0.02522	0.02500	0.02451
	1	0.02628	0.02376	0.02371	0.02354	0.02430
	2	0.02435	0.02368	0.02355	0.23531	0.02348
	3	0.02416	*0.02350*	0.02351	0.02381	0.02344
	4	0.02381	0.02351	0.02348	0.02345	0.02344

AIC

		MA				
		0	1	2	3	4
AR	0	-1141.56	-1198.53	-1294.5	-1306.62	-1336.6
	1	-1231.21	-1386.95	-1389.37	-1399.35	-1396.2
	2	-1348.85	-1390.78	-1398.75	-1398.85	-1401.04
	3	-1360.1	*-1402.08*	-1400.2	-1381.86	-1402.6
	4	-1381.86	-1400.09	-1401.04	-1401.95	-1402.09

SIC

		MA				
		0	1	2	3	4
AR	0	-1082.74	-1134.36	-1224.98	-1231.76	-1256.59
	1	-1167.04	-1317.43	-1314.51	-1319.14	-1319.4
	2	-1279.33	-1315.92	-1318.54	-1313.29	-1310.13
	3	-1285.24	*-1321.88*	-1316.2	-1301.65	-1306.35
	4	-1301.65	-1314.54	-1310.13	-1305.7	-1300.49

Chi-square statistic p-value to lag 6

		MA				
		0	1	2	3	4
AR	0	0	0	0	0	0
	1	0	0.0004	0.0088	0.09	0.0124
	2	0	0.0246	0.5339	0.3397	0.3226
	3	0	*0.133*	0.4511	0.9505	0.8975
	4	0.0344	0.04	0.3226	0.4055	0.6485

Table F.3: Relevant sample statistics in the selection of an appropriate ARMA process for the residuals of the Dynamic Regression model of the natural log of release

F.5 Final Dynamic Regression model of the natural log of release

Variable	Slope estimate	Standard error	Approx t-value	P-value
Constant	4.38925	0.11672	37.61	<.0001
MA 1 term	0.89564	0.02504	35.77	<.0001
AR 1 term	1.01231	0.03675	27.55	<.0001
AR 2 term	0.11151	0.03616	3.08	0.0021
AR 3 term	-0.1442	0.03029	-4.76	<.0001
EST_WV	-0.02318	0.002268	-10.22	<.0001
PRICE	0.000583	0.000163	3.58	0.0004
INFLOW_L2	0.05424	0.01929	2.81	0.005
INFLOW_L1	-0.18342	0.02753	-6.66	<.0001
INFLOW	0.40288	0.02954	13.64	<.0001
INFLOW_F1	0.18295	0.02959	6.18	<.0001
INFLOW_F2	-0.44266	0.02758	-16.05	<.0001
INFLOW_F3	0.12479	0.01893	6.59	<.0001
SAT	-0.07849	0.01043	-7.53	<.0001
SUN	-0.08617	0.01058	-8.15	<.0001
Variance	0.02350			
AIC	-1402.08			
SBC	-1321.88			

Table F.4: Final parameter estimates, approximate t-values and p-values for the Dynamic Regression model of logged release

As mentioned above, the residuals of the final model, the estimated parameter values of which are listed in Table F.4, must be checked to ensure there are no patterns and there is no serial correlation present in the model residuals. None of the Chi-square statistics listed below in Table F.5 are statistically significant at the 5% level, and there are no significant sample autocorrelation or partial autocorrelation coefficients at the 5% level (as shown in Figure F.3 and Figure F.4).

F.6 Chi-square statistics for model residuals

To Lag:	Chi-Square	DF	Pr > ChiSq
6	4.03	2	0.133
12	7.2	8	0.5149
18	9.79	14	0.7773
24	14.93	20	0.7804
30	19.6	26	0.8102
36	26.94	32	0.7208
42	28.52	38	0.8677
48	30.76	44	0.9347

Table F.5: Chi-square statistics and p-values for the residuals of the final Dynamic Regression model, for lags in multiples of six

F.7 Autocorrelogram and partial autocorrelogram of final Dynamic Regression model residuals

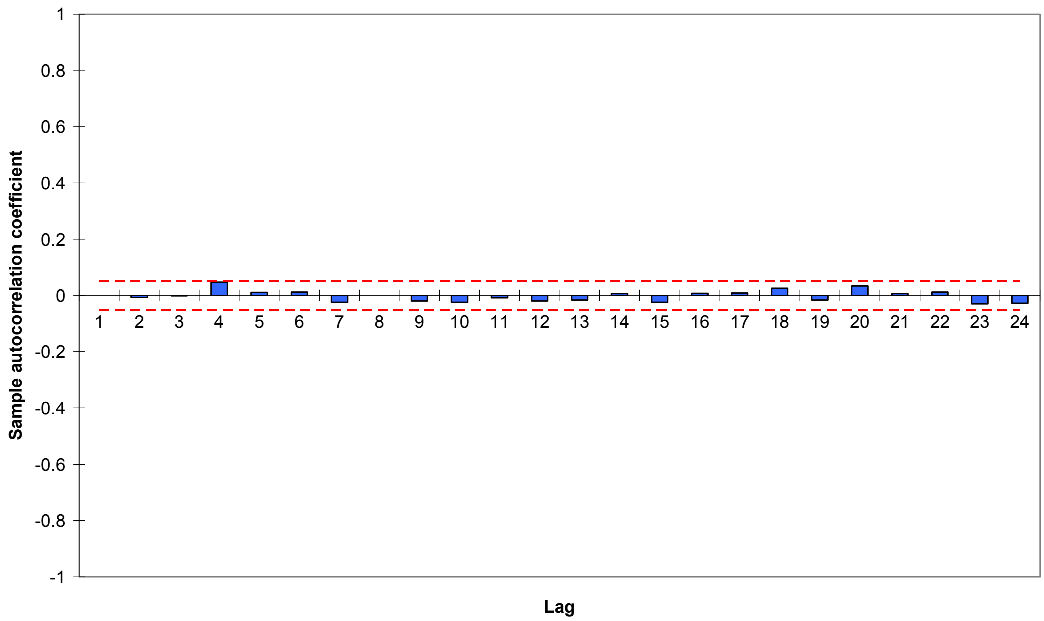

Figure F.3: Autocorrelogram of the residuals of the final DR equation listed in Table F.4

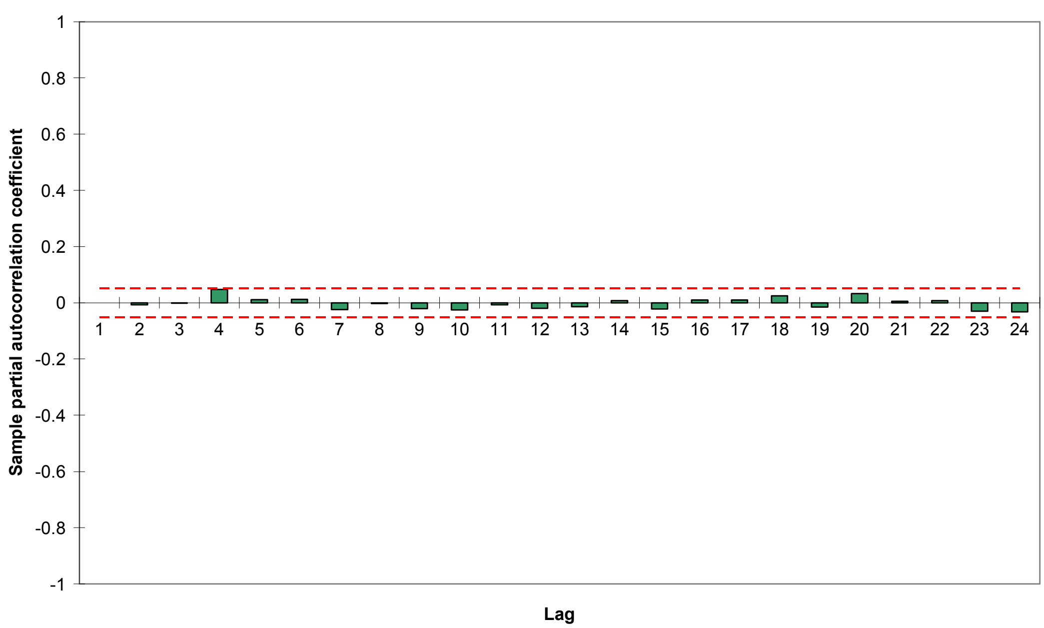

Figure F.4: Partial autocorrelogram of the residuals of the final DR equation listed in Table F.4

Appendix G

TESTING DIFFERENCES IN NZEM STORAGE POLICY

This appendix contains tests of hypotheses regarding the difference between storage regimes in New Zealand since 1980, as described in Chapter 7.

Two different tests are applied repeatedly throughout this appendix. The first of these tests whether or not the mean value of a single distribution is statistically different to zero. This is useful for determining, for example, whether or not the true mean difference between the average annual simulated storage level and the average annual observed storage level is zero, in which case it could be concluded that there is no difference between the simulated and actual results, or whether the difference is different from zero. A one-tailed test can conclude whether or not the mean of a distribution is greater (or less) than zero, whereas a two-tailed test just concludes whether or not the mean is different to zero. This test calculates a test statistic, z, based on the mean value of the sample distribution and the standard error of the sample distribution (which is based on the estimated variance and number of observations in the sample). The z-statistic for sample i is calculated with the following formula:

$$z_i = \frac{\bar{x}_i}{s_i / \sqrt{n_i}}$$

where $\bar{x}_i$ is the sample mean of the distribution, s_i is the sample standard deviation, and n_i is the sample size. If the sample size is larger than 30, which it is in all cases in this appendix, the test statistic can then be compared with the standard normal distribution, regardless of the underlying distribution of the observations. This test standardizes the sample mean and finds the probability (or p-value) of drawing a standardized sample mean of that magnitude from a standard normal distribution. If the p-value is less than a chosen threshold (say, 5%) then the null hypothesis is rejected; the standardized sample mean is concluded to be significantly different from zero.

The second test used in this appendix tests whether the mean values of two different distributions are statistically significantly different from each other. A one-tailed test can conclude that the mean of one distribution is significantly greater (or less) than the mean of another distribution, whereas a two-tailed test concludes simply whether or not the two mean values are different from each other. This test also requires the calculation of a z-statistic using the sample means, standard deviations and sizes of the two sample distributions. The z-statistic to test whether or not the mean of sample i is significantly different from the mean of sample j is shown below:

$$z_{i-j} = \frac{\bar{x}_i - \bar{x}_j}{\sqrt{\dfrac{s^2{}_i}{n_i} + \dfrac{s^2{}_j}{n_j}}}$$

With large sample sizes (i.e. both samples > 30), the z-statistic can again be compared with the standard normal distribution to calculate a p-value. If the p-value is less than a chosen threshold (say, 5%) then the null hypothesis is rejected; the difference between the standardized sample means is concluded to be significantly different from zero.

G.1 Testing the differences between annual average storage levels

The sample statistics for each of the three distributions in Section 7.3.1 of Chapter 7 are:

$$n_{MoE} = 1400 \qquad \bar{x}_{MoE} = 110.91 \qquad s^2_{MoE} = 65916.89$$
$$n_{ECNZ} = 2000 \qquad \bar{x}_{ECNZ} = -111.78 \qquad s^2_{ECNZ} = 69819.65$$
$$n_{NZEM} = 1400 \qquad \bar{x}_{NZEM} = -7.16 \qquad s^2_{NZEM} = 65144.66$$

Firstly, it is worthwhile testing the null hypothesis that the true mean simulated difference between observed annual average storage levels and simulated annual average storage levels in the market era is not significantly different from zero, to ensure that the simulation over the period 1997-2003 produces appropriate results.

$$H_0: \quad \mu_{NZEM} = 0$$
$$H_a: \quad \mu_{NZEM} \neq 0 \qquad \text{(two-tailed test)}$$

The test statistic for this test is calculated as described above:

$$z_{NZEM} = \frac{\bar{x}_{NZEM}}{s_{NZEM} / \sqrt{n_{NZEM}}} = \frac{-7.16}{255.23 / \sqrt{1400}} = -1.05$$

The p-value associated with a test statistic value of -1.05 is 27.93%, which is greater than the rejection value of 5%. Therefore, there is a 27.93% chance that the true population mean difference in the NZEM era is zero given the sample data observed; this is too great a probability to reject the null hypothesis. We can make the inference therefore that the annual average storage levels simulated by the model using the inflows from 1997 to

2003 are no different from the annual average storage levels observed between 1997 and 2003.

Similar hypothesis tests for the mean differences in the MOE and ECNZ eras yield *z*-statistics of 16.16 and -18.92, which both have p-values of 0.0000% (to four decimal places). Therefore we have enough evidence to reject the null hypothesis in both cases, and conclude that the annual average storage levels simulated by the model in the MOE (ECNZ) years using the inflows from 1980 to 1986 (1987-1996) are significantly different from the annual average storage levels observed between 1980 and 1986 (1987-1996).

Further hypothesis testing using the second type of test can be undertaken to determine whether the mean difference from the MoE era is different to that of the NZEM era, and likewise to compare the mean differences of the ECNZ and NZEM eras. Firstly, it appears as though the true mean of all differences in the MoE era is likely to be greater than the true mean of all differences in the NZEM era, therefore that is the alternative hypothesis tested:

H_0: $\mu_{MoE} - \mu_{NZEM} = 0$

H_a: $\mu_{MoE} - \mu_{NZEM} > 0$ (one-tailed test)

The test statistic for this test is calculated as described above:

$$z_{MoE-NZEM} = \frac{\bar{x}_{MoE} - \bar{x}_{NZEM}}{\sqrt{\frac{s^2_{MoE}}{n_{MoE}} + \frac{s^2_{NZEM}}{n_{NZEM}}}} = \frac{110.91 + 7.16}{\sqrt{\frac{65916.89}{1400} + \frac{65144.66}{1400}}} = 12.20$$

For the null hypothesis not to be rejected, the probability of obtaining a test statistic with a value at least as large as 12.20 from a standard normal distribution must be greater than 5%. However, the probability of drawing a value of 12.20 from a standard normal distribution is 0.0000% (to four decimal places). Therefore, we have enough evidence at the 5% level to reject the null hypothesis, and conclude that the true mean difference between observed storage levels and simulated levels is greater in the MoE era than in the NZEM era.

Repeating the procedure to test whether or not the true mean difference in the ECNZ era is less than that of the NZEM era yields a test statistic of 11.59 (again with a p-value of 0.0000%). Therefore, we have enough evidence at the 5% level to reject the null hypothesis, and conclude that the true mean difference between observed storage levels and simulated levels is greater for the NZEM era than for the ECNZ era.

G.2 Testing the differences between annual minimum storage levels

The sample statistics for each of the three distributions in Section 7.3.2 of Chapter 7 are:

$$n_{MoE} = 1400 \qquad \bar{x}_{MoE} = 171.01 \qquad s^2_{MoE} = 182794.54$$

$$n_{ECNZ} = 2000 \qquad \bar{x}_{ECNZ} = -24.61 \qquad s^2_{ECNZ} = 154555.56$$

$$n_{NZEM} = 1400 \qquad \bar{x}_{NZEM} = 118.49 \qquad s^2_{NZEM} = 127463.19$$

Tests of the null hypotheses that each population mean equals zero reject the null hypothesis in all three cases, with z-statistics (p-values) of 14.97 (0.0000%) for the MOE era, -2.80 (0.5872%) for ECNZ and 12.42 (0.0000%) for the NZEM. Similarly, tests reject the null hypotheses that the population mean of the NZEM era equals that of the MOE era (z-statistic = 3.53), and that the population mean of the NZEM era equals that of the ECNZ era (z-statistic = 11.03).

G.3 Testing the differences between annual EOY storage levels

The sample statistics for each of the three distributions in Section 7.3.3 of Chapter 7 are:

n_{MoE}	=	1400	$\bar{x}_{MoE}$	=	166.16	s^2_{MoE}	=	169899.43
n_{ECNZ}	=	2000	$\bar{x}_{ECNZ}$	=	-173.68	s^2_{ECNZ}	=	264550.66
n_{NZEM}	=	1400	$\bar{x}_{NZEM}$	=	127.14	s^2_{NZEM}	=	186076.16

Tests of the null hypotheses that each population mean equals zero reject the null hypothesis in all three cases, with z-statistics (p-values) of 15.08 (0.0000%) for the MOE era, -15.10 (0.0000%) for ECNZ and 11.03 (0.0000%) for the NZEM. Similarly, tests reject the null hypotheses that the population mean of the NZEM era equals that of the MOE era (z-statistic = 2.45), and that the population mean of the NZEM era equals that of the ECNZ era (z-statistic = 18.47).

Appendix H

VERIFICATION OF THE T-CONE MODEL

H.1 Introduction

This appendix contains analysis of the generation and price output from repeated runs of the T-CONE model used in Chapter 8. Each run varies in the combination of the pseudo elasticity of demand (PED) of the four regions and the pseudo contract level (PCL) of the generating companies (see Chapter 8 for descriptions of these two variables). The results are explained in the context of theoretical results for Cournot models.

H.2 Input data

The input data for T-CONE that is unchanged in each run includes information on the generators, such as their location, capacity, SRMC and the company they are owned by, and also on the transmission constraints between the regions (including the Snowy Region). All generating companies are assumed to be Cournot players. The factors that are varied from run to run are the demand curve for each state and the PCL of each company, which is specified as a percentage of that company's least-cost dispatch, as explained in Chapter 8. In our runs, each region (QLD, NSW, VIC and SA) has the same PED (Snowy has no load), and each generating company has the same PCL.

H.3 Demand curve calculation

While the runs of the model undertaken in Chapter 8 cover 731 days of discrete data, from 1 January 2003 – 31 December 2004, all illustrations in this document are the result of running the model using the data from 9 October 2003 only. This enables a more thorough examination of the effects of changing the PCL and PED.

For each day the model is run, T-CONE requires a load/price pairing through which the demand curve is drawn in each region. The loads used are a weighted average of the actual load observed for each day[1], while prices are paired to each load as follows. If the load observed in VIC on 9 October 2003 is the 100th largest out of 731, then it is matched with the 100th highest price in VIC out of the 731 observed over the length of the sample period. This results in a monotonically non-decreasing price/load relationship over the course of the 731 days.

Given a load/price pairing for each of the four regions, and a single value for the elasticity of demand across each of the four regions, four linear demand curves can be drawn in the form $Price = \alpha - \beta \times Quantity$, as shown in Figure H.1 below. Each of the coloured dots corresponds to a region's load/price pairing, and the slope of the line drawn through that pairing is determined by the elasticity.

[1] Each of the 48 half-hourly loads from each day was aggregated into one of three periods for that day: peak, shoulder and off-peak, and then averaged within each period to give a peak load, a shoulder load and an off-peak load for each day of the sample (see Chapter 8).

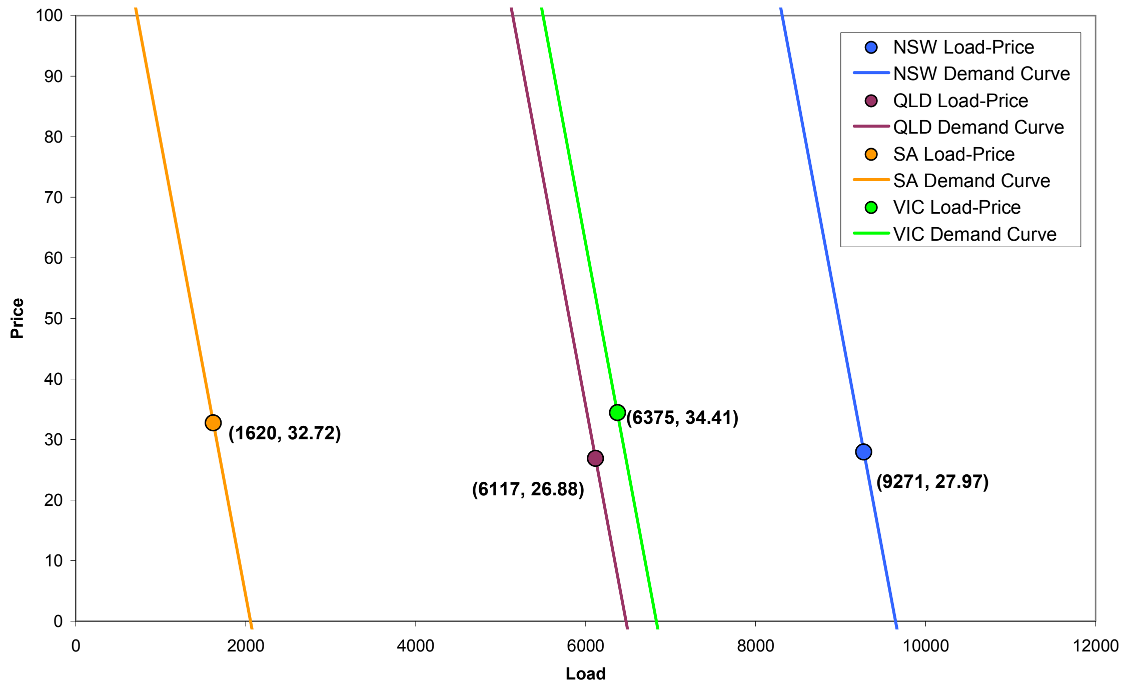

Figure H.1: Conceptual demand curves for 9 October 2003

H.4 Model results

The model was solved repeatedly for these four load/price pairings, varying both the PED and the PCL, in order to assess the effects on the price and total generation. For the purposes of this document, the PED was varied from -0.001 to -15 (-0.001, -0.01, -0.05, -0.15, -0.5, -1, -1.5, -2, -5, -15), and the PCL from 0% to 110% (0%, 25%, 50%, 75%, 90%, 100%, 110%).

The prices and aggregate market generation levels resulting from these runs, plotted on price and quantity axes, are shown in Figure H.2. Holding the PED (and hence the demand curve) constant and varying the PCL, then plotting lines between the solutions, effectively plots the demand curve for each level of the PED. Note that all the demand curves intersect at a common price/load point (shown by the black dot), as expected. This point remains fixed for each demand curve, while the slope of the curve changes and the curves pivot around that point.

As expected, the demand curves become steeper as the PED approaches zero, and flatter as it increases in magnitude. The leftmost and highest point on each demand curve is given by the 0% PCL (yielding the highest price and lowest generation for a given elasticity) and the rightmost and lowest point on each curve is given by the 110% PCL. This result is consistent with the theory of Cournot models – increasing the contract level for a given demand curve increases generation and decreases the spot price.

With the vertical axis on the graph being truncated, you cannot see the high prices which result from very inelastic demand (the red and green lines). Note that for a PED of -0.5 (light blue line), the majority of the demand curve is to the right of the common load/price point, with the points for 75%, 90%, 100% and 110% PCL all being well to the right of the 0% point, and very close together. They appear almost to be converging on a single point.

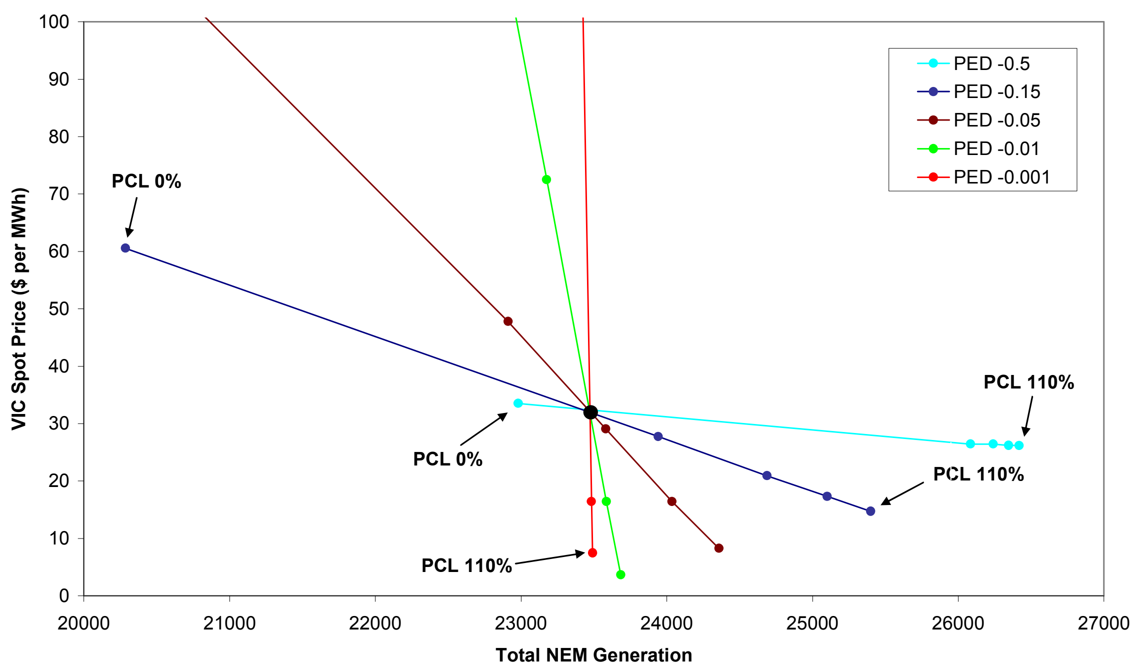

Figure H.2: Aggregate generation versus VIC spot prices while varying PED between -0.5 and -0.001 and PCL between 0% and 110%. Each solid line is formed by holding PED constant and varying the PCL

Looking in greater detail at the right-hand end of the -0.5 PED demand curve, and adding more demand curves for higher levels of the PED, gives more of an indication what is occurring as the PED increases. Figure H.3 shows not only the -0.5 PED demand curve but also curves generated by PED of -1, -1.5, -2, -5 and -15. The pink line (a PED of -1.5) shows that the range of generation and prices between a PCL of 0% and 110% is much smaller with such a high PED. The higher the PED, the less a part the PCL plays in determining the price and generation level, until for a PED of -5 or greater, when the PCL is *completely* insignificant. The generation level converges on a single MW amount, for a single price, regardless of the PCL. At this point the level of demand is so responsive to a change of price that generating companies are unable to exercise any market power at all, and contracting is not necessary. They will basically produce the same amount (26,950 MW) regardless of the price they receive for that generation. But note that the price is higher for higher PED – *not* because more market power is exercised, but because of the way the assumed demand curve swings around the observed market point, until it is completely horizontal.

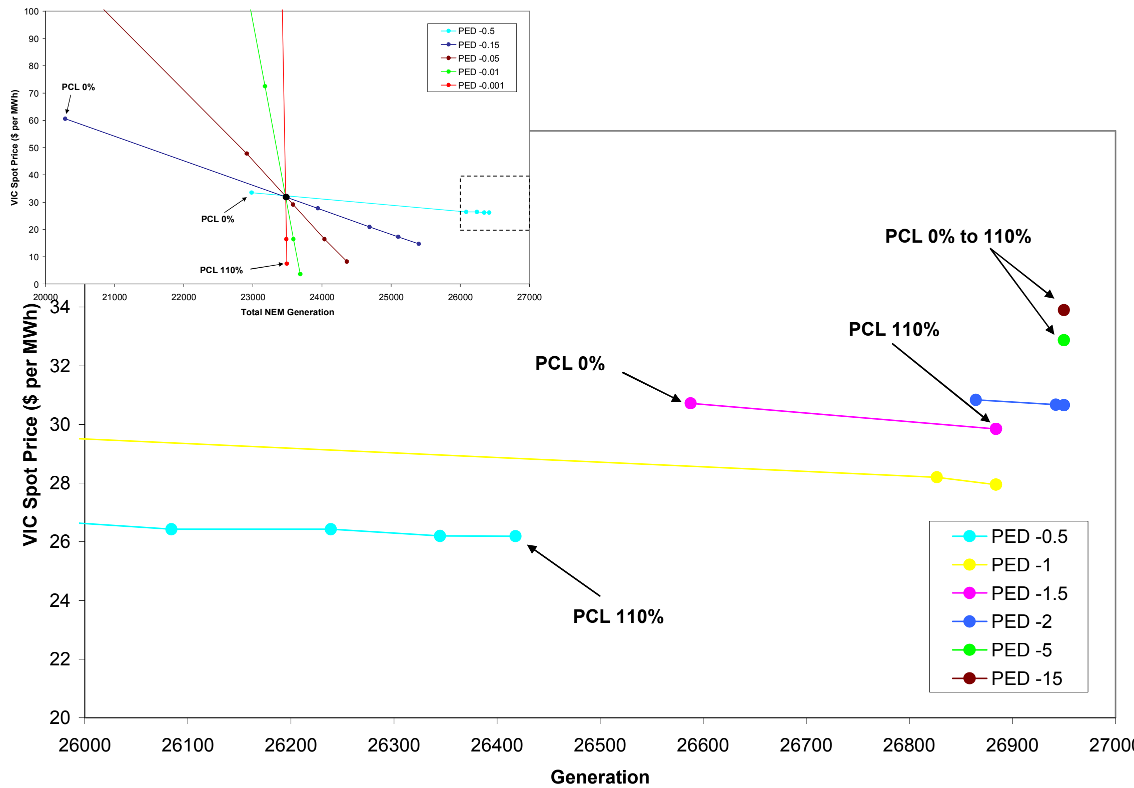

Figure H.3: Aggregate generation versus VIC spot prices while varying PED between -0.5 and -15, and PCL between 0% and 110%

Figure H.4 shows the effects on the VIC spot price of varying the PED while holding various PCL constant, and reveals some slightly anomalous behaviour. The initial movement in prices is downwards for all PCL as the PED decreases from -2 (moving from left to right), which is contrary to what you would expect. With more inelastic demand, prices (somewhat surprisingly) continue to fall for PCL of 100% and over 100%, but eventually begin to rise for PCL under 100%, which is what you would expect.

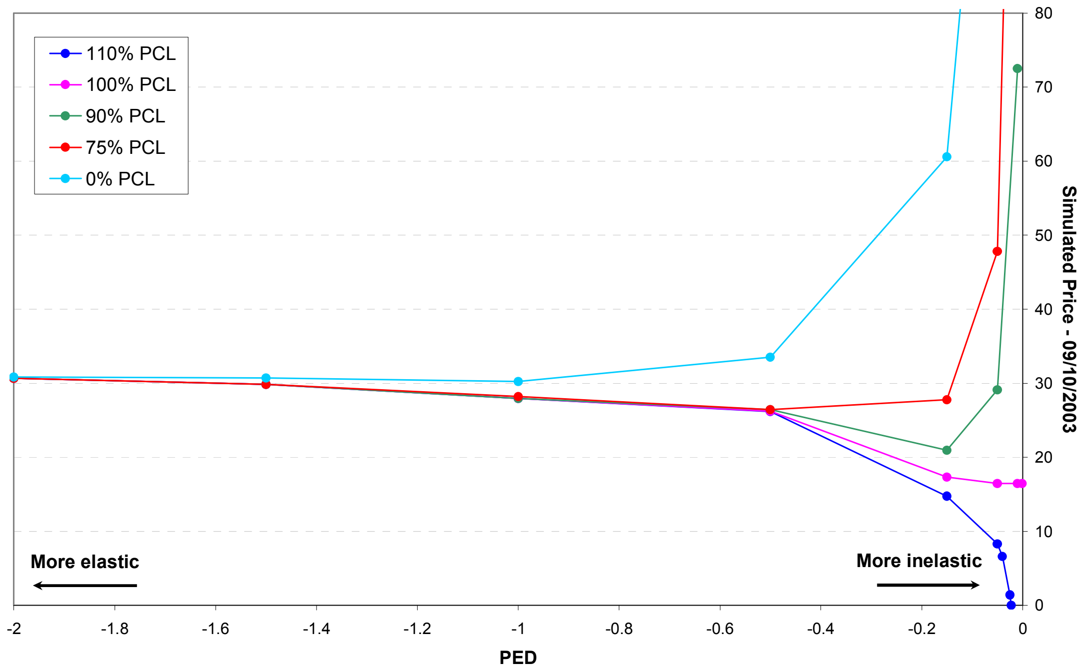

Figure H.4: VIC spot prices versus PED while holding the PCL constant.

The problem with the curves plotted Figure H.4 is not that curves are diverging as the PED decreases towards 0, or the direction of divergence, as these in line with expected behaviour. What is anomalous is the fact that as the PED decreases, the pink 100% PCL line is not horizontal, and prices do not simply increase as elasticity decreases. This is because the demand curves swing around a certain point depending on the PED, giving a different perfectly competitive solution for each of the individual PED. The perfectly competitive solution for each PED lies where its particular demand curve intersects the aggregate SRMC curve, therefore varying the PED will trace out the SRMC curve at these intersections. As the demand curve becomes steeper, the intersection occurs at a point lower down the SRMC curve than before, giving a lower price, and resulting in the pink line shown in Figure H.4.

This concept is better illustrated graphically. Connecting up the dots in Figure H.2 in another way yields the levels of generation offered given different PCL. As mentioned in the previous paragraph, varying the PED but holding the PCL constant at 100% should trace out the perfectly competitive supply curve, as (theoretically) generating companies have no incentive to increase prices above their marginal cost. This is shown in Figure H.5. Every point on the 100% curve (the black line) should correspond to the intersection of a particular demand curve with the SRMC curve. As the SRMC curve is monotonically non-decreasing, so is the 100% curve. The 110% contract level curve is lower than the 100%, as expected, and the curves for the lower PCL are higher.

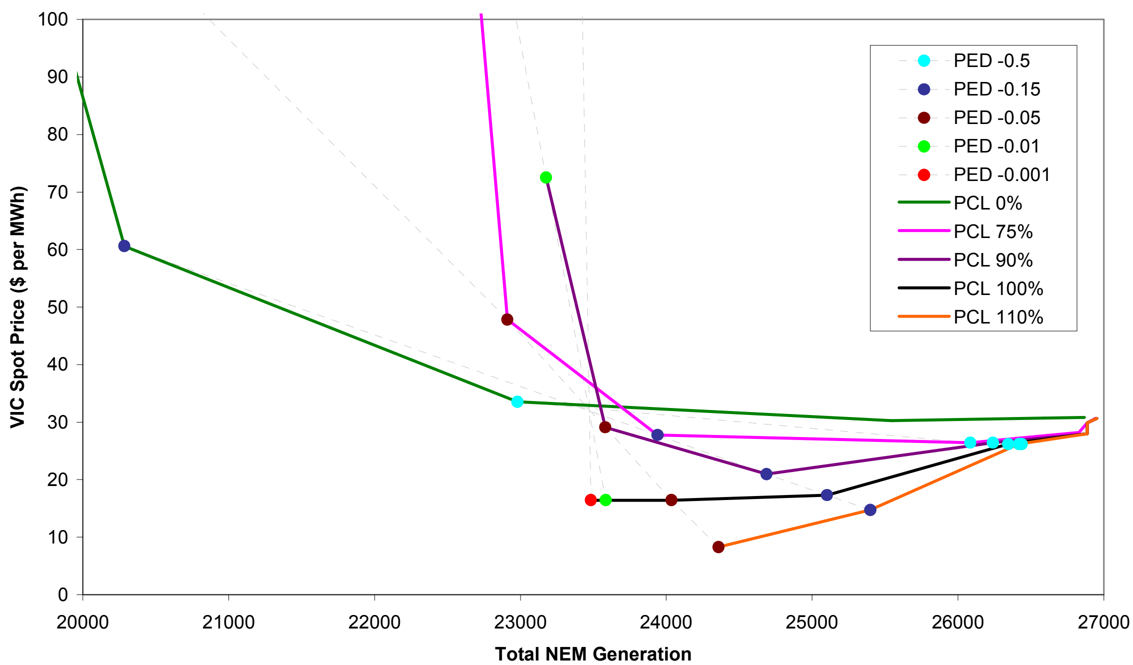

Figure H.5: Aggregate generation versus VIC spot prices while varying the PED between -0.5 and -0.001 and the PCL between 0% and 110%. Each solid line is formed by holding the PCL constant and varying the PED. The dashed lines correspond to the demand curves shown in Figure H.2.

Whereas the curves in Figure H.5 for the 100% and 110% PCL slope down as the PED becomes less elastic, you might expect the curves corresponding to PCL less than 100% only to increase above the perfectly competitive price as the PED becomes more elastic. Intuitively, generation should decrease and prices should increase as the PED decreases. However, each of the curves (including the 0% PCL) slopes up, finally, toward the very high elasticity points at the middle right of Figure H.3, but for less elastic demand the curves slope down (before some of them increase). This is an unexpected result, as is the fact that some of these curves cross over each other.

There are two separate factors at work influencing the price and generation in these cases. Firstly, as the PED becomes less elastic (moving from right to left in Figure H.5), total generation decreases and dispatched generators with the highest SRMCs will be taken offline, lowering the spot price[2]. Secondly, as demand becomes less elastic, generating companies with a PCL lower than 100% have an incentive to decrease generation and raise their offer prices, which increases the spot price. This second effect obviously dominates the first once the PED is more inelastic than -0.15, as prices begin to rise for even the 90% PCL. This price effect is also shown in Figure H.4. As the PED becomes more inelastic (i.e. moving from left to right in Figure H.4), prices decrease for higher levels of the PED and only begin to increase towards the right hand side of the graph. The green line of the 90% PCL on that graph gives a good illustration of the two effects vying with each other, one pulling the price down and the other pulling it up.

Figure H.5 also showed that aggregate generation increases as demand becomes more elastic, which, as mentioned earlier, is the expected result. Figure H.6 plots aggregate generation versus PED, for various PCL, confirming this result. As we move from left to right on the graph, decreasing elasticity, generation levels decrease for all PCL.

[2] Linked with this effect is the fact that as generation increases, transmission constraints are more likely to become binding, which will cause prices on the capacity-constrained side of the constraint to increase and prices on the other side to decrease. A decrease in generation reverses this effect.

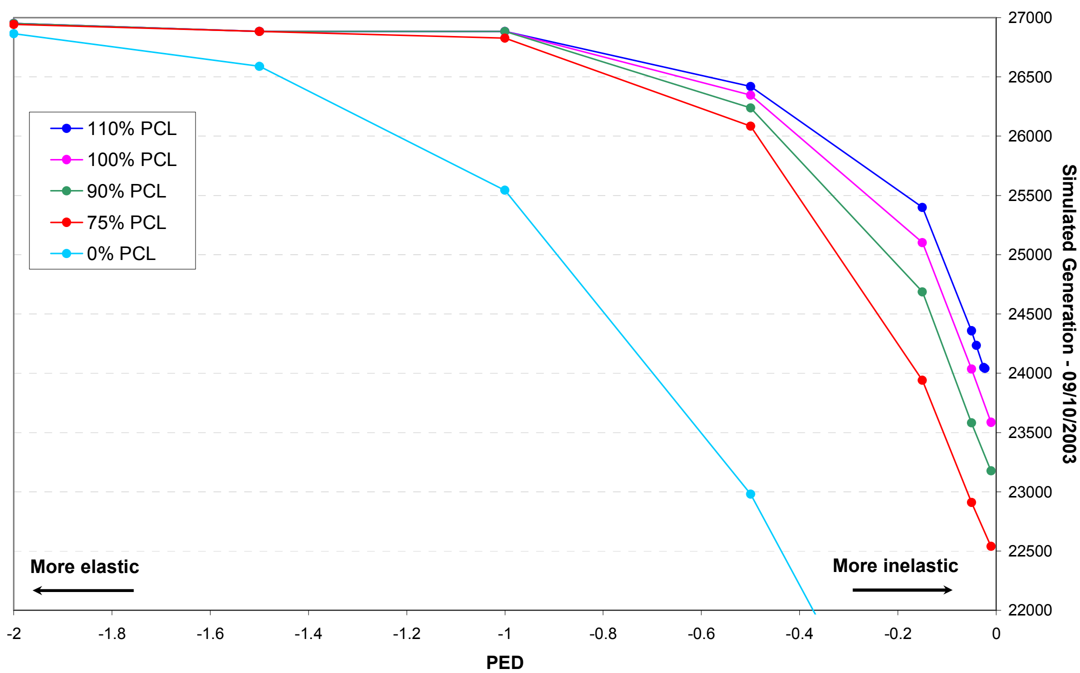

Figure H.6: Aggregate generation versus PED while holding the PCL constant.

H.5 Conclusion

While an initial examination of the results may suggest otherwise, the model is in fact performing in line with the established theory on Cournot models. The fact that prices do not necessarily increase as the PED becomes more inelastic is an interesting result, and a worthwhile discovery in its own right.

Appendix I

NEM PRICE MODEL ESTIMATION

This appendix contains the estimated parameters of the stochastic price processes fitted to Australian peak, shoulder and off-peak prices in 2003-4 in Chapter 8. Table I.1 contains the estimated parameters from the full EPV model for each of the three series. Table I.2 contains the estimated stochastic process parameters for the three series of residuals, and Table I.3 contains the estimated parameters from the full EPV model for just the series of peak residuals.

I.1 Estimation of full EPV model to peak, shoulder and off-peak prices

	Peak			**Shoulder**			**Off-peak**	
	Coeff.	t-stat		Coeff.	t-stat		Coeff.	t-stat
CONSTANT	26.38	5.87		13.20	118.32		10.78	10.92
TREND	0.031	6.83		0.017	0.00		0.010	7.58
M_2	-7.64	-1.51		-0.42	-13.16		-2.30	-2.24

M_3	-10.71	-2.20		-1.74	-43.63		-2.06	-1.90
M_4	-5.51	-1.15		-0.38	-10.40		-0.96	-0.84
M_5	7.05	1.45		2.70	58.34		1.07	0.95
M_6	10.42	1.96		0.26	8.22		0.30	0.24
M_7	-0.36	-0.07		1.31	30.78		0.33	0.29
M_8	-8.51	-1.76		-1.55	-37.80		-1.21	-1.06
M_9	-9.10	-1.87		1.64	51.89		0.57	0.50
M_{10}	-8.34	-1.68		-0.10	-3.69		0.41	0.32
M_{11}	-11.10	-2.27		2.46	71.59		0.36	0.30
M_{12}	-11.37	-2.34		4.14	119.54		-0.62	-0.43
WEEKDAY	4.59	5.02		4.46	31.39		2.15	11.61
λ	4.16%	4.76		10.53%	0.04		10.83%	2.23
θ	0.48	14.55		0.60	0.39		0.62	20.05
ω	83.38	10.57		5.83	87.82		0.55	2.14
α	0.35	4.54		0.53	0.94		0.11	3.77
β	0.00	0.62		0.00	0.00		0.73	10.15
μ	330.38	2.55		4.84	324.54		-1.24	-1.31
σ^2	409855.4	3.48		319.87	328455.2		27.68	2.90
Log-Likelihood	-3093			-2360			-1736	
SIC	6324			4859			3610	

Table I.1: Estimated parameters for the fitted EPV model for each of the peak, shoulder and off-peak series of Victorian spot prices, 1 January 2003 – 31 December 2004

I.2 Estimation of EPV stochastic component to peak, shoulder and off-peak residuals

	Peak			**Shoulder**			**Off-peak**	
	Coeff.	t-stat		Coeff.	t-stat		Coeff.	t-stat
λ	4.30%	5.15		4.32%	4.43		6.11%	2.37
θ	0.53	16.72		0.77	30.56		0.77	29.57
ω	89.31	9.01		14.90	11.70		0.54	2.86
α	0.20	4.47		0.29	5.58		0.12	4.07
β	0.11	3.14		0.00	0.13		0.74	13.68
μ	325.04	2.86		29.71	2.89		1.29	1.09
σ^2	325588.3	4.27		2094.91	3.06		39.20	2.47
Log-Likelihood	-3185			-2341			-1731	
SIC	6508			4821			3601	

Table I.2: Estimated parameters for the fitted EPV stochastic component for each of the peak, shoulder and off-peak residuals, 1 January 2003 – 31 December 2004

I.3 Estimation of full EPV model to peak residuals

	Peak	
	Coeff.	t-stat
CONSTANT	2.21	0.78
TREND	0.019	5.78
M_2	-3.19	-0.95
M_3	-3.77	-1.17
M_4	3.44	1.05
M_5	9.48	2.87
M_6	4.77	1.36
M_7	-13.64	-3.95
M_8	-14.36	-4.39
M_9	-1.77	-0.54
M_{10}	1.22	0.35
M_{11}	-0.08	-0.02
M_{12}	-3.42	-1.02
WEEKDAY	-4.74	-5.06
λ	4.44%	5.37
θ	0.34	12.32
ω	84.66	11.69
α	0.19	4.34
β	0.01	0.98
μ	317.83	2.75
σ^2	384918.3	3.80
Log-Likelihood	-3062	
SIC	6261	

Table I.3: Estimated parameters for the fitted EPV model to the peak residuals, 1 January 2003 – 31 December 2004

Appendix J

DIRECTIONS FOR FURTHER RESEARCH

The research presented in this thesis has both filled in and extended the academic literature. However, in the course of achieving these goals, many other questions have been raised and several natural extensions have been identified. A number of these are described below, with some explored thoroughly and others still at the conceptual stage.

J.1 Modelling higher frequency prices or including volatility in the MWV calculation

The decision to model daily average prices in this research, as opposed to prices with a higher frequency, was made partly because of constraints in computer processing speed. The CML routine in the GAUSS modelling package was unable to process a series 48 (or more) times as long as the daily time series we use in this study, and therefore the extension to modelling intra-day volatility using half-hourly final spot prices was not done.

Figure J.1 below plots the Waitaki RSL versus half-hourly Benmore spot prices, and illustrates that the same exponential-type relationship between the RSL and the spot price is evident even when half-hourly prices replace daily average prices. However, there are many periods in which the storage level was relatively very high, and the prices still peaked at levels around \$500-\$600/MWh. This suggests that the price spikes were a result of some factor other than a shortage of water in the major long-term storage reservoirs. The analysis of historic offer stacks would go some way to determining why this price behaviour occurred.

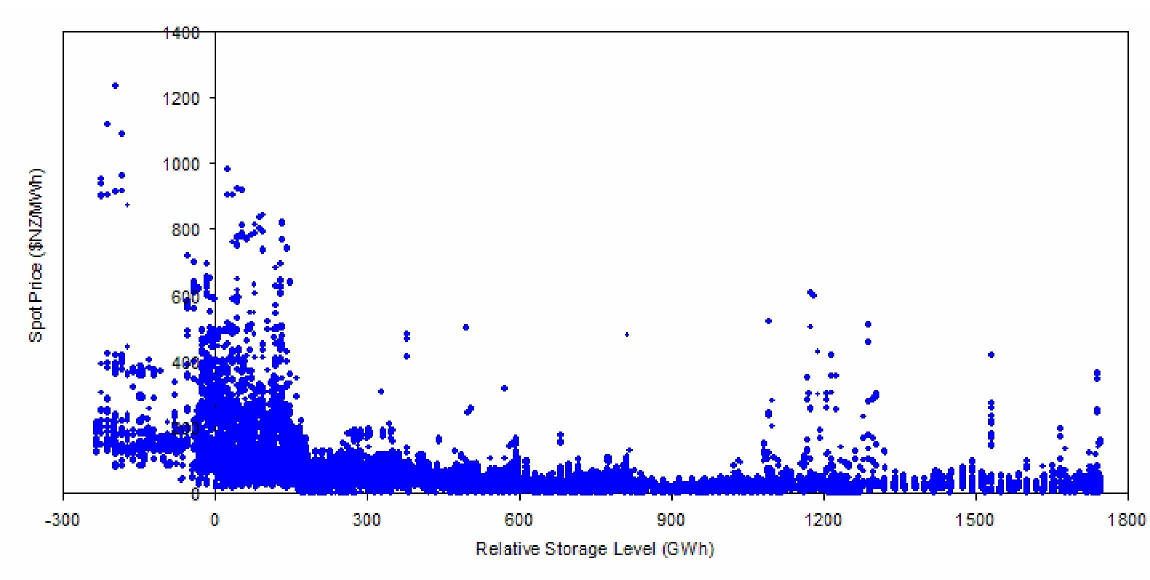

Figure J.1: Half-hourly final spot prices from the Benmore node versus daily Waitaki RSL, August 1999 - June 2003

As mentioned in Chapter 3, Guthrie and Videbeck (2002b) study the behaviour of high-frequency prices in the NZEM, and conclude that prices can be grouped into five separate intra-day markets on the basis of the correlations between prices at certain times of the day. In Chapter 5 it was illustrated that intra-day volatility (as represented by the intra-day standard deviation of prices) is strongly linked to the RSL. Therefore, combining the RSL methodology with Guthrie and Videbeck's intra-day market approach would appear a worthwhile extension to the current NZEM price model. However, any high-frequency analysis involving storage levels would need also to look at the storage and release behaviour at individual, smaller reservoirs around the country.

Guthrie (personal communication, 2007) also suggests that because the MWV used in reservoir management can be interpreted as "an option value from retaining water for future use" (see Chapter 4), option-pricing theory suggests that the MWV will be an increasing function of spot price volatility. The estimated function for the MWV used in this thesis is a function only of the RSL, however he suggests it could be enhanced by including some measure of spot price volatility in the function. While it is not clear how this may be achieved, "an alternative might be to estimate the standard deviation of daily price changes over the preceding month, and let that feed into the MWV calculation."

J.2 Using a non-symmetric distribution for the stochastic component

Many studies, including Knittel and Roberts (2005) and Escribano et al. (2002), suggest that Gaussian models cannot account for very large changes in price. As represented by the New Zealand price time series, while there are some negative jumps, price spikes are predominately positive in nature, therefore an extension to exponentially-distributed jumps (as in Villaplana, 2003) may be advantageous. Both Deng, Jiang and Zhendong (2002) and Natarajan (2003) found that incorporating error distributions with fatter tails than the normal distribution accounted for an increased proportion of volatility. The move away from Gaussian models and towards more asymmetric processes, especially with regards to the jump process, would be a very worthwhile step to take[1].

[1] Furthermore, the GARCH process could be replaced with an asymmetric volatility process, such as exponential GARCH (EGARCH), as used by Knittel and Roberts (2005). These authors noted that electricity prices exhibit greater volatility after positive shocks than after negative shocks, which they called the "inverse leverage effect", and used the EGARCH process to model this behaviour. Duffie et al. (1999) also use an EGARCH process to model spot price volatility.

J.3 Adding other physical information

Besides the hydro storage level, other related physical information could be included in the underlying model for NZEM prices and releases. Due to the recent public availability of time series of NZEM load, this would the most obvious extra piece of information to include, and may improve the fitting performance of the model. However, as mentioned in Chapter 2, load is often included as an explanatory variable in price models, but is not always statistically significant. It is noteworthy that our model performs as well as it does without load.

When this research was undertaken, we considered using the air temperature as a proxy for load. Knittel and Roberts (2005) found that temperature variables were statistically significant in their models for Californian spot price (except during the 2001 crisis period). Due to the long, thin shape of New Zealand, the temperature varies quite markedly across the country at any one time. However, when the entire country is caught in a cold snap at once, load can increase dramatically, as occurred on 17 August 2004 when load hit record levels in the North Island, and near record levels in the South (New Zealand Press Association, NZPA, 2004). However, the extent to which the load influences the spot price is highly dependent on the storage level at the time. At that time in 2004, while the load was extremely high the RSL was also very high, and thus the effect on prices was not as great as it would have been had the RSL been lower.

Including data on generating plant and transmission outages, as well as flows across the inter-island HVDC link, would also be worthwhile. As illustrated in Chapter 8, when base-load plant is taken offline it can have a significant impact on prices. Similarly, when transmission links fail or reach their capacity, markets can become segmented, seriously impacting the prices in the different segments.

J.4 A piecewise linear model for the MWV

The RSL methodology used throughout this thesis proved to be particularly adept at identifying the MWV from storage information; however, it is not the only method that could be used.

In some pre-market era work on scheduling releases from hydro reservoirs, Boshier, Manning and Read (1983) and Read and Boshier (1989) use an interesting and intuitive method for calculating New Zealand's overall MWV. They calculate storage level contours for the whole year, with each contour representing a level when specific non-hydro generation units would need to be dispatched to reduce the risk of running out of water. The rationale behind this method is similar to the RSL methodology. Assuming generating capacity is large enough, when the RSL is very high, the entire country's load can be met with hydro generation. When the RSL drops below a certain level (represented by a contour), the cheapest non-hydro unit will need to be dispatched to ensure long-term security of supply. When it drops further, below another contour, the next cheapest unit must be dispatched, and so on. Using this method, the exact MWV can be calculated when the storage level is on one of contours – it is exactly the SRMC of the unit corresponding to that contour. In between the contours, the MWV can be calculated by interpolation.

We experimented with using a variation of this method to calculate the MWV for storage levels between 1999 and 2003[2]. Instead of being calculated exactly using dynamic programming, each contour is a straight line that slopes upwards for the part of the year when storage is increasing, peaks when storage is usually at its maximum level (around the start of April), and slopes downwards until the time when storage is at its usual minimum level (around November). Once a set of contours has been defined, each can be assigned a value. The highest contours (i.e. those corresponding to the greatest storage levels) have the lowest MWV, and the lowest have the greatest MWV.

[2] All analysis in this section should be considered preliminary, and was undertaken by hand, without any formal estimation.

The method is best explained graphically. Figure J.2 below shows a set of four piecewise linear contours, each corresponding to a different MWV, and the observed aggregate NZEM storage level of 2001. Note that the contours in Figure J.2 at the end of October are much more compact than at the end of March. This is due to the seasonal nature of New Zealand's load and inflows – a high storage level is required in autumn, before inflows decrease and loads decrease, to ensure security of supply throughout the winter. In contrast, a high storage level at the end of winter may result in some of the spring inflows having to be spilled.

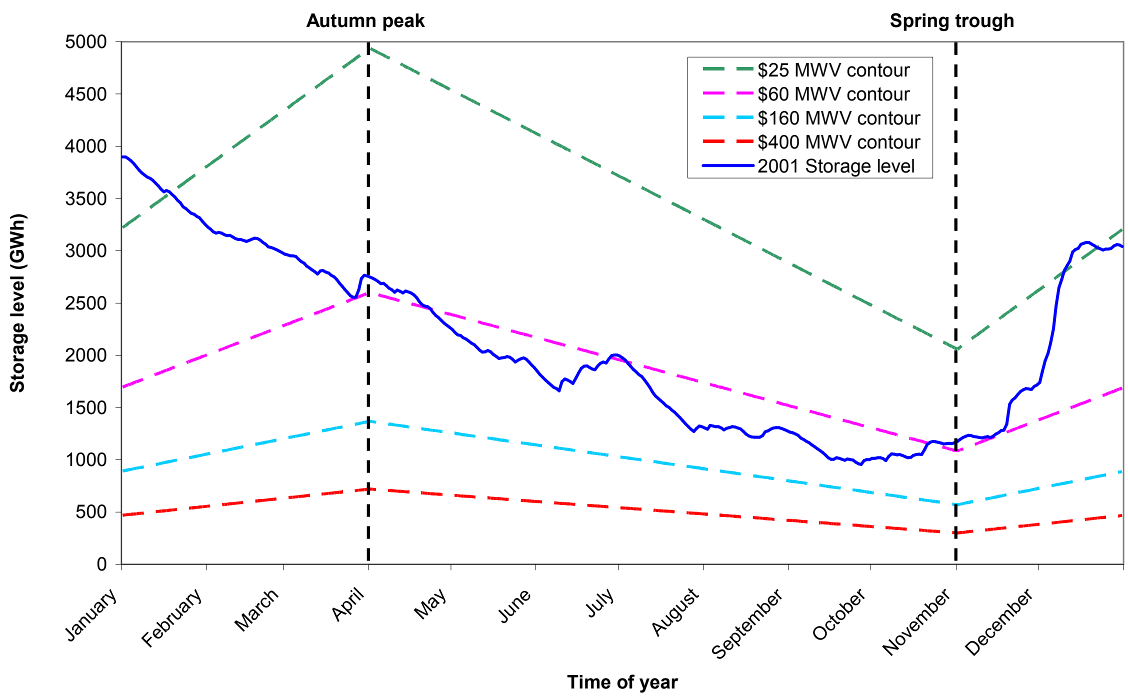

Figure J.2: Annual linear MWV contours for the aggregate NZEM storage level, with the observed storage trajectory from 2001 superimposed

At the start of 2001, the storage level was high, and, as it is above the $25/MWh contour, it would have been assigned a MWV between $0 (the MWV corresponding to having to spill water) and $25. Near the end of January, the storage level crossed over the $25/MWh contour, which, if this was the model of Boshier et al, would have required the cheapest non-hydro unit (that had a SRMC of $25/MWh) to be dispatched to reduce the risk of a future water shortage. From February to April, the MWV would have been between \$25/MWh and \$60/MWh, with its exact value calculated using linear interpolation. In mid-April, the storage trajectory crossed over the \$60/MWh contour (meaning the non-hydro unit with an SRMC of \$60/MWh would need to be dispatched), and for most of the time from then until mid-October the MWV would be linearly interpolated between \$60/MWh and \$160/MWh. This MWV path can be seen in Figure J.4.

The MWV curves corresponding to the range of possible storage levels at the time of the autumn peak and spring trough can be plotted, as shown in Figure J.3. These two curves are used to calculate the estimated MWV at any time during the year. For example, if the observed storage level was 1200 GWh, the estimated MWV would be around \$200/MWh if it were the 1st of April, but only around \$55/MWh if it were the 1st of November. If the storage level were 1250GWh on 1 September, as it was in 2001, then the linearly interpolated MWV would be approximately \$200 - 5/7 × (\$200-\$55) ≈ \$96/MWh.

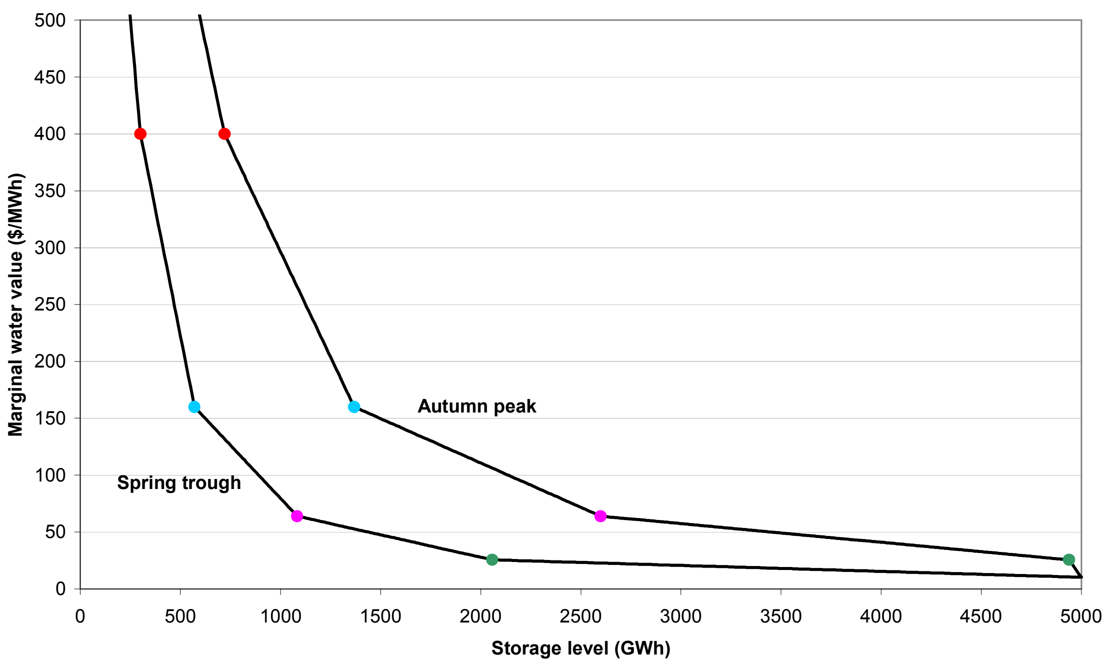

Figure J.3: MWV curves corresponding to the spring trough and autumn peak in Figure J.2

Like the RSL method used throughout this thesis, using the method of linear MWV contours and linearly-interpolated MWVs can produce a very good fit to observed market prices. Figure J.4 shows the daily average spot prices from the Haywards node from 1999-2003, and the estimated MWVs using this method; it is clear that the estimated MWVs track the underlying price level very well.

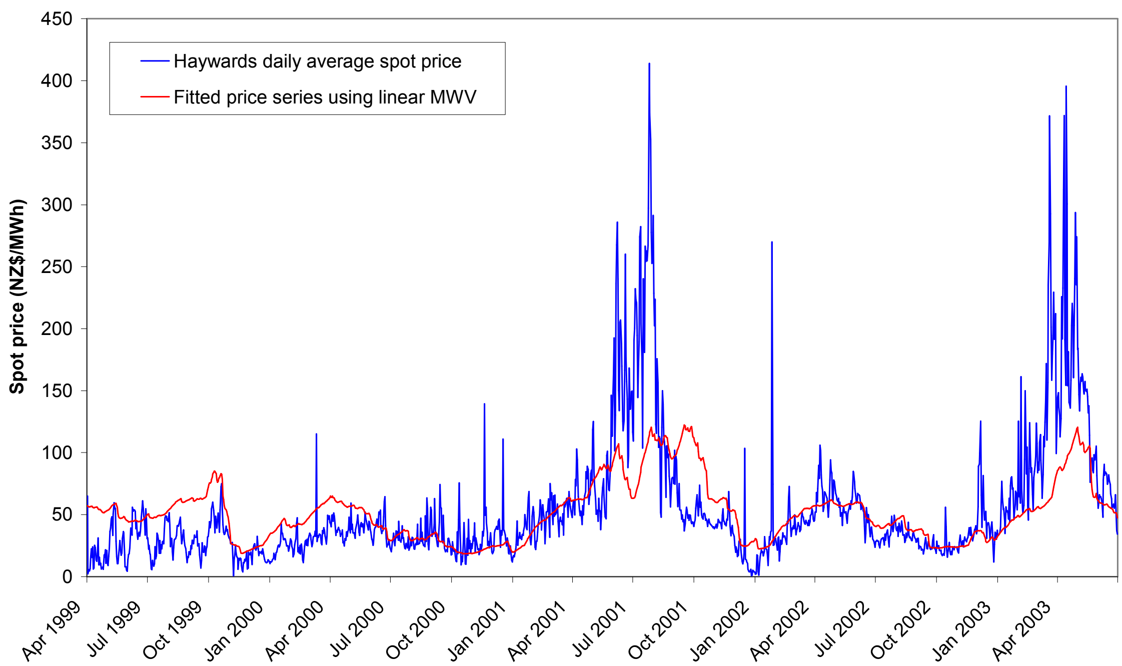

Figure J.4: Daily average spot prices from the Haywards node, April 1999 – June 2003 and modelled prices using the linear MWV contours

Potentially, this method requires a minimum of only six parameters to be estimated: the timing of the annual peak and trough, the number of contours, the minimum MWV at the peak, the ratio between the prices on each contour and the ratio between the prices at the peak and trough. Complexity could be added by allowing more breakpoints in the contours between the peak and trough.

J.5 A two-island price and release model for the NZEM

The geography of the NZEM gives it some interesting operating characteristics that differ from many other markets. Traditionally in New Zealand, the majority of generating capacity has been in the lower half of the South Island (SI) and the majority of demand in the upper half of the North Island (NI). Between the two islands is a High Voltage Direct Current cable, with limited transmission capacity. Therefore even if the SI hydro lakes are completely full, when transmission to the NI is constrained by the link's capacity, there may be major differences between nodal spot prices in the two islands.

As electricity can be transferred between the islands, using a two-reservoir model would be superior to the current simulation model presented, which uses just one. For example, if storage were low in the NI but plentiful in the South, spot prices would still be low over the whole country provided inter-island transfers did not exceed the capacity of the link. Similarly, if NI storage levels were much higher than SI storage, transfers across the link from north to south would ensure that the prices were still low, and roughly equivalent, throughout the country. However, as soon as flows across the link become constrained, the prices in the island with surplus generating capacity (usually, but not restricted to, the SI) decrease, while more expensive generation needs to be called upon in the other island, raising spot prices there.

This means that each island should have its own water value function, which would be a function of both NI and SI storage levels[3]. For example, if storage is relatively plentiful in the SI then its MWV would be low, provided only the SI's storage level is considered in the calculation. However, if, at the same time, relative storage were low in the NI, then the SI's MWV would be increased to reflect that relative shortage. Even if SI storage

[3] The approach detailed in this section is consistent with that of Read, Culy, Halliburton and Winter (1988) in their PRISM simulation model for electricity planning in New Zealand. PRISM aggregates storage into two reservoirs, NI and SI. A two-dimensional grid of storage levels is constructed, which shows release priorities for each island given the respective levels of the two reservoirs. The priorities are determined by the two storage levels (through the calculation of water values), HVDC transfer capacity, and the costs of thermal generation in the NI. The location of the current storage pair is found on the grid, and the amount of release from each reservoir can then be calculated.

were completely full, spot prices would still not be zero if storage in the North were low. Similarly, the NI water value would be very high if its calculation were based on just the storage in that island. However, including the relative surplus in the SI in that calculation would reduce the MWV to reflect the fact that the overall national storage situation is not as severe as the North's RSL would suggest. Given a release function in the same form as in the current model, release would therefore be much greater in the SI, where the MWV was lower, than in the North.

Either from these modelled releases, or the calculated water values, some inference could be made about the likely inter-island transfers. If the MWV were near zero in the SI but very high in the North, there would obviously be a reasonably high probability of the link capacity being fully utilised through transfers northwards. A study examining the water values and releases in each island at times when the inter-island transfers were constrained would yield an appropriate model for the probabilities and price effects of such events.

The steps required in such a study would be:

1. Calculate the RSL for the NI and SI.
2. Estimate parameters for the two functions: $Price_{NI} = f(RSL_{NI}, RSL_{SI})$ and $Price_{SI} = f(RSL_{SI}, RSL_{NI})$ to get the MWV curves for each island.
3. Calculate the series of releases for each island, and then fit release models to both.
4. Analyse the effects that HVDC transmission (and in particular constrained transmission) have on prices.

Having capacity-constrained links between regions is, in itself, not unique to New Zealand, and an extended model of this kind would be of use to other markets as well. For example, in Australia, Tasmania is about to be connected to the national network, and 90% of its electricity comes from hydro generation. The prices in Victoria and Tasmania and the differences between these prices will depend very heavily on the cost and availability of hydro generation in Tasmania (which has a maximum total generating capacity of over 2000MW) and the capacity of the link (630MW) between the two regions. In other overseas markets with a high proportion of hydro generation and limited transfer capacity[4], the same factors will influence regional prices and inter-regional price differences.

J.6 Using the MWV calibration methodology in a Cournot model

The two major pieces of research presented in this thesis could be combined into a single Cournot model for NZEM prices, using the RSL methodology to determine the MWV of the hydro generating companies.

After obtaining a series of NZEM load, and the SRMC for each of the non-hydro generating plants, a first step in this analysis would be to extract the MWV function for each island using market data and the steps detailed above in Section J.5[5]. Heuristics could then be used to determine the MWV for each of the hydro generating companies – depending on the level of detail required and the purpose of the calibration, this may require more explicit modelling of the hydro storage levels than simply using two reservoirs. Then, once all the required input data had been collected, the Cournot calibration could be undertaken, using a similar methodology to that used for the Australian market study in Chapter 8[6].

---

[4] Another example, on a much larger scale, is the hydropower station located at Foz do Iguaçu on the border between Paraguay and Brazil. The station has a generating capacity of over 12000MW, but only 6300MW can be injected directly onto the Brazilian grid. Another 6000MW can be transported straight to the industrial area of São Paulo via the Itaipu HVDC link.

[5] Alternatively, Wolak (2003b) proposes a method for extracting individual firms' cost functions from observed bid data, which is essentially what this step of the methodology is required to achieve. To date, however, it is not apparent whether or not Wolak has applied his method successfully to extracting the cost functions of hydro-generating companies.

[6] Modifications to the methodology should be made, for example taking into account prolonged generator outages and also using the calibrated Cournot price as a driver for the stochastic process.

As was calculated for the NEM, the generation level of each NZEM market participant would be calculated in each period using the available SRMC (and MWV), with the generation levels determining approximately how much water would be released from each reservoir. Instead of simply calibrating the Cournot model by matching the NZEM prices, the model would have to be able to model releases accurately and hence be able to calculate storage trajectories and MWV for each reservoir. This would ensure that the storage trajectories predicted by the Cournot model matched reality, which could be achieved by maximising the fit to both the underlying price level and the storage levels. The top-down model for NZEM releases would therefore not be used to calculate releases; however, it could be used as a cross-check. Any price and storage forecasts made using the calibrated Cournot model could then be cross-checked against hydro simulation model presented in this thesis, giving two separate forecasting procedures.

Once the compilation and calibration of the NZEM model had been completed, there are many studies for which it would be of use. For example, it could examine the effects and cost-effectiveness of increasing the capacity of the inter-island link or adding a new link between the islands, increasing transmission capacity to the upper NI, or splitting the four major generating companies into smaller, competing companies. Also, the effects of increasing the degree to which each generating company is contracted could be assessed. In the current New Zealand market environment, this would be a useful tool.